

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 3, 2006, 10:53:23 ; Search time 863 Seconds
(without alignments)
1646.682 Million cell updates/sec

Title: US-10-719-900-1

Perfect score: 25
Sequence: 1 aaaaaaacacgcgcagcatgg 25

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_in.*

3: gb_env.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pt.*

9: gb_to.*

10: gb_sts.*

11: gb_sy.*

12: gb_un.*

13: gb_vi.*

14: gb_mtg.*

15: gb_pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	25	100.0	221399	9	AL604045	AL604045 Mouse DNA
2	23.4	93.6	49499	8	AF184110	AF184110 Homo sapi
3	23.4	93.6	51905	6	AR659626	AR659626 Sequence
4	23.4	93.6	51905	6	AR659740	AR659740 Sequence
5	23.4	93.6	61412	14	AC021493	AC021493 Homo sapi
6	23.4	93.6	139280	8	AC099052	AC099052 Homo sapi
7	23.4	93.6	142943	8	AC099405	AC099405 Papio anu
8	23.4	93.6	150221	8	AC004593	AC004593 Homo sapi
9	23.4	93.6	163473	14	AC144870	AC144870 Pan trogl
10	23.4	93.6	165802	8	AL135914	AL135914 Human DNA
11	23.4	93.6	168813	8	AP000555	AP000555 Homo sapi
12	23.4	93.6	171272	8	AC146225	AC146225 Pan trogl
13	23.4	93.6	171461	8	AC092041	AC092041 Homo sapi
14	23.4	93.6	184515	8	AC010653	AC010653 Homo sapi
15	23.4	93.6	185161	8	AC006059	AC006059 Homo sapi
16	23.4	93.6	187530	14	AC096875	AC096875 Pan trogl
17	23.4	93.6	199471	8	AC009097	AC009097 Homo sapi
18	23.4	93.6	328439	14	AL590137	AL590137 Homo sapi

AB030316	Mus muscu	9	AB030316	12593	89.6	22.4	C 19
BV02806	RPANMSEQO	10	BV02806	394	87.2	21.8	C 20
BV09496	RPANMSEQO	10	BV09496	398	87.2	21.8	C 21
BV159481	RPANMSEQO	10	BV159481	398	87.2	21.8	C 22
AF270524	Homo sapi	8	AF270524	701	87.2	21.8	C 23
BV669872	S21686216	10	BV669872	749	87.2	21.8	C 24
BV590677	G5916435	10	BV590677	861	87.2	21.8	C 25
BV594288	bxj38a02.	10	BV594288	942	87.2	21.8	C 26
AK090515	Homo sapi	8	AK090515	2141	87.2	21.8	C 27
AK097083	Homo sapi	8	AK097083	2530	87.2	21.8	C 28
S81092	acyl-coenzy	9	S81092	3041	87.2	21.8	C 29
L42293	Mus musculu	9	MUSACACT	3697	87.2	21.8	C 30
BS571793	Human DNA	8	BS571793	35616	87.2	21.8	C 31
AC104230	Homo sapi	14	AC104230	46784	87.2	21.8	C 32
AL358393	Human DNA	8	AL358393	50175	87.2	21.8	C 33
AC104800	Homo sapi	8	AC104800	54618	87.2	21.8	C 34
AC105044	Homo sapi	14	AC105044	65882	87.2	21.8	C 35
AL589910	Human DNA	8	AL589910	66591	87.2	21.8	C 36
AC133141	Homo sapi	8	AC133141	70560	87.2	21.8	C 37
AL591766	Human DNA	8	AL591766	73084	87.2	21.8	C 38
AP000280	Homo sapi	8	AP000280	73687	87.2	21.8	C 39
AC026374	Homo sapi	8	AC026374	75385	87.2	21.8	C 40
AL096775	Human DNA	8	HSDJ336M4	76241	87.2	21.8	C 41
AP000039	Homo sapi	8	AP000039	100000	87.2	21.8	C 42
AP000107	Homo sapi	8	AP000107	100000	87.2	21.8	C 43
AP000183	Homo sapi	8	AP000183	100000	87.2	21.8	C 44
Continuation (2 of		14	AC145312_1	110000	87.2	21.8	C 45
AC012356	Homo sapi	8	AC012356	112527	87.2	21.8	C 46
AL109811	Human DNA	8	HSJ635E18	112769	87.2	21.8	C 47
AL008721	Human DNA	8	HSJ390C10	114231	87.2	21.8	C 48
AC078901	Homo sapi	14	AC078901	115909	87.2	21.8	C 49
AL121820	Human chr	8	CNS01DSM	117449	87.2	21.8	C 50
AP000609	Homo sapi	8	AP000609	130964	87.2	21.8	C 51
AC008131	Homo sapi	8	AC008131	131891	87.2	21.8	C 52
AC006001	Homo sapi	8	AC006001	135044	87.2	21.8	C 53
AL954211	Pan trogl	8	AL954211	138573	87.2	21.8	C 54
AC103718	Homo sapi	8	AC103718	144801	87.2	21.8	C 55
AC146516	Homo sapi	14	AC146516	147478	87.2	21.8	C 56
AJ300188	Homo sapi	8	HSA300188	148049	87.2	21.8	C 57
AC044819	Homo sapi	14	AC044819	151798	87.2	21.8	C 58
AC011921	Homo sapi	8	AC011921	153922	87.2	21.8	C 59
AC148077	Pan trogl	14	AC148077	155690	87.2	21.8	C 60
AC092198	Homo sapi	8	AC092198	158456	87.2	21.8	C 61
AC166354	Mus muscu	14	AC166354	160029	87.2	21.8	C 62
AC120013	Mus muscu	14	AC120013	161646	87.2	21.8	C 63
AC087434	Pan trogl	8	AC087434	161716	87.2	21.8	C 64
AC145756	Pan trogl	14	AC145756	162240	87.2	21.8	C 65
AC023421	Homo sapi	8	AC023421	163365	87.2	21.8	C 66
AC109460	Homo sapi	8	AC109460	163424	87.2	21.8	C 67
AL391259	Human DNA	8	AL391259	163520	87.2	21.8	C 68
AC010823	Homo sapi	14	AC010823	164376	87.2	21.8	C 69
AC068023	Homo sapi	14	AC068023	164383	87.2	21.8	C 70
AC161221	Mus muscu	14	AC161221	164547	87.2	21.8	C 71
AC067860	Homo sapi	14	AC067860	165264	87.2	21.8	C 72
AL45467	Human DNA	8	AL45467	166336	87.2	21.8	C 73
AL4591722	Homo sapi	14	AL4591722	166439	87.2	21.8	C 74
AC1356864	Homo sapi	14	AC1356864	166876	87.2	21.8	C 75
AC101652	Mus muscu	9	AC101652	168183	87.2	21.8	C 76
AC079047	Homo sapi	14	AC079047	169372	87.2	21.8	C 77
AC138627	Homo sapi	8	AC138627	170920	87.2	21.8	C 78
AC016178	Homo sapi	8	AC016178	171035	87.2	21.8	C 79
AL355517	Human DNA	8	AL355517	171348	87.2	21.8	C 80
AC069433	Homo sapi	14	AC069433	171544	87.2	21.8	C 81
AL160396	Human DNA	8	AL160396	172437	87.2	21.8	C 82
AC140275	Mus muscu	9	AC140275	173022	87.2	21.8	C 83
AL590292	Homo sapi	14	AL590292	173751	87.2	21.8	C 84
AF216798	Homo sapi	14	AF216798	174295	87.2	21.8	C 85
AC135506	Homo sapi	8	AC135506	174363	87.2	21.8	C 86
AC069125	Homo sapi	14	AC069125	176010	87.2	21.8	C 87
AL353807	Human DNA	8	AL353807	179064	87.2	21.8	C 88
AC114486	Homo sapi	8	AC114486	179070	87.2	21.8	C 89
AC149106	Papio anu	14	AC149106	179560	87.2	21.8	C 90
AL355840	Human chr	8	CNS05TCV	179937	87.2	21.8	C 91

92	21.8	87.2	181842	9	AC137127	AC137127 Mus muscu	ORGANISM	Mus musculus
93	21.8	87.2	183483	14	AC136510	AC136510 Pan trogl	REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
94	21.8	87.2	183539	8	AC092268	AC092268 Homo sapi	AUTHORS	Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
95	21.8	87.2	184000	8	AP002449	AP002449 Homo sapi	TITLE	Sciuognathi; Muroidea; Muridae; Murinae; Mus.
96	21.8	87.2	184295	14	AC021138	AC021138 Homo sapi	JOURNAL	1
97	21.8	87.2	185146	14	AC104311	AC104311 Homo sapi		Whitehead,S.
98	21.8	87.2	186134	9	AL591390	AL591390 Mouse DNA	Direct Submission	Submitted (04-APR-2002) Wellcome Trust Sanger Institute, Hinxton,
99	21.8	87.2	187788	14	AC138925	AC138925 Homo sapi		Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
100	21.8	87.2	191137	14	AC145064	AC145064 Pan trogl	COMMENT	humquy@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk
101	21.8	87.2	191196	14	AL441903	AL441903 Homo sapi		On Apr 7, 2002 this sequence version replaced gi:17127811.
102	21.8	87.2	191957	8	AC015845	AC015845 Homo sapi		During sequence assembly data is compared from overlapping clones.
103	21.8	87.2	194160	8	AC130472	AC130472 Homo sapi		Where differences are found these are annotated as variations
104	21.8	87.2	194164	14	AC146119	AC146119 Pan trogl		together with a note of the overlapping clone name. Note that the
105	21.8	87.2	195558	8	AC007448	AC007448 Homo sapi		variation annotation may not be found in the sequence submission
106	21.8	87.2	196954	8	AC008267	AC008267 Homo sapi		corresponding to the overlapping clone, as we submit sequences with
107	21.8	87.2	197683	9	AC118476	AC118476 Mus muscu		only a small overlap as described above.
108	21.8	87.2	200028	8	AC092680	AC092680 Homo sapi		This sequence was finished as follows unless otherwise noted: all
109	21.8	87.2	201020	8	AC022809	AC022809 Homo sapi		regions were either double-stranded or sequenced with an alternate
110	21.8	87.2	201709	14	AC145497	AC145497 Papio anu		chemistry or covered by high quality data (i.e., phred quality >=
111	21.8	87.2	207860	14	AP001907	AP001907 Homo sapi		30); an attempt was made to resolve all sequencing problems, such
112	21.8	87.2	209472	14	AC090759	AC090759 Homo sapi		as compressions and repeats; all regions were covered by at least
113	21.8	87.2	214019	14	AC064803	AC064803 Mus muscu		one plasmid subclone or more than one M13 subclone; and the
114	21.8	87.2	214331	14	AC145462	AC145462 Callithri		assembly was confirmed by restriction digest. The following
115	21.8	87.2	219646	14	AC139247	AC139247 Homo sapi		abbreviations are used to associate primary accession numbers given
116	21.8	87.2	219646	14	AC139247	AC139247 Homo sapi		in the feature table with their source databases: Em; EMBL; Swi;
117	21.8	87.2	228363	9	AC116128	AC116128 Mus muscu		SWISSPROT; Tr; TREMBL; Wp; WORMPEP; Information on the WORMPEP
118	21.8	87.2	232188	9	AL672055	AL672055 Mouse DNA		database can be found at
119	21.8	87.2	232401	14	AC106137	AC106137 Rattus no		http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-418011 is
120	21.8	87.2	240058	14	AC110921	AC110921 Homo sapi		from the RPI-23 Mouse PAC Library
121	21.8	87.2	247451	14	AC097417	AC097417 Rattus no		constructed by the group of Pieter de Jong.
122	21.8	87.2	248281	8	AC008737	AC008737 Homo sapi		For further details see http://www.chori.org/bacpac/home.htm
123	21.8	87.2	256019	8	AC016525	AC016525 Homo sapi	FEATURES	VECTOR: pBACE3.6.
124	21.8	87.2	259474	8	HUAC004605	HUAC004605 Homo sapi	source	Location/Qualifiers
125	21.8	87.2	340000	8	AP001715	AP001715 Homo sapi		1. .221399
126	21.4	85.6	215965	14	AC151869	AC151869 Callithri		/organism="Mus musculus"
127	21	84.0	204793	8	AC026333	AC026333 Homo sapi		/mol_type="genomic DNA"
128	20.8	83.2	401	10	BV190162	BV190162 sqmml6629		/db_xref="taxon:10090"
129	20.8	83.2	55673	8	AC117944	AC117944 Homo sapi		/chromosome="11"
130	20.8	83.2	96322	8	AL512640	AL512640 Human DNA		/clone="RP23-418011"
131	20.8	83.2	97832	8	HSJ46801	HSJ46801 Ruman DNA		/clone_lib="RPCI-23"
132	20.8	83.2	101324	8	AC010386	AC010386 Homo sapi	ORIGIN	
133	20.8	83.2	104947	8	AC008116	AC008116 Homo sapi		Query Match 100.0%; Score 25; DB 9; Length 221399;
134	20.8	83.2	107469	8	HS29H4	AL078475 Homo sapi		Best Local Similarity 100.0%; Pred.No. 0.4; Indels 0; Gaps 0;
135	20.8	83.2	112895	8	AC107210	AC107210 Homo sapi		Matches 25; Conservative 0; Mismatches 0;
136	20.8	83.2	117879	8	AC110004	AC110004 Homo sapi		
137	20.8	83.2	129090	14	AC083961	AC083961 Homo sapi	QY	1 AAAAAAAAAATCAGCGCCAGCATGG 25
138	20.8	83.2	132070	8	AC003663	AC003663 Homo sapi	Db	AAAAAAAAATCAGCGCCAGCATGG 21812
139	20.8	83.2	146947	14	AC012515	AC012515 Homo sapi		
140	20.8	83.2	150810	8	AC018500	AC018500 Homo sapi		
141	20.8	83.2	152333	8	AC010476	AC010476 Homo sapi		
142	20.8	83.2	157356	8	AL590783	AL590783 Human DNA		
143	20.8	83.2	160298	14	AC067818	AC067818 Homo sapi	RESULT 2	AF184110/c
144	20.8	83.2	161139	8	AC132812	AC132812 Homo sapi	LOCUS	AF184110
145	20.8	83.2	162364	8	AP005213	AP005213 Homo sapi	DEFINITION	Homo sapiens cyclophilin-related protein (NKTR) gene, complete cds.
146	20.8	83.2	163489	8	AC016697	AC016697 Homo sapi	ACCESSION	AF184110
147	20.8	83.2	164888	14	AC024101	AC024101 Homo sapi	VERSION	AF184110.1 GI:5923890
148	20.8	83.2	169199	8	AC026463	AC026463 Homo sapi	KEYWORDS	.
149	20.8	83.2	176600	14	AC023546	AC023546 Homo sapi	SOURCE	Homo sapiens (human)
150	20.8	83.2	178050	8	AC142324	AC142324 Pan trogl	ORGANISM	Homo sapiens
								Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
								Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
								Hominidae; Homo.
								1 (bases 1 to 49499)
								Anderson,S.K., Gallinger,S., Roder,J., Frey,J., Young,H.A. and
								Ortaldo,J.R.
								A cyclophilin-related protein involved in the function of natural
								killer cells
								Proc. Natl. Acad. Sci. U.S.A. 90 (2), 542-546 (1993)
								8421688
								2 (bases 1 to 49499)
								Anderson,S.K.

ALIGNMENTS

RESULT 1	AL604045	221399 bp	DNA	linear	ROD 05-APR-2002
LOCUS	AL604045				
DEFINITION	Mouse DNA sequence from clone RP23-418011 on chromosome 11,				
	complete sequence.				
ACCESSION	AL604045				
VERSION	AL604045.7	GI:20068530			
KEYWORDS	HTG.				
SOURCE	Mus musculus (house mouse)				


```

ORIGIN
Query Match          93.6%; Score 23.4; DB 8; Length 49499;
Best Local Similarity 96.0%; Pred.No.2.5;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 AAAAAAAAAATCAGCGCCAGGCATGG 25
         |||||
Db       16079 AAAAAAAAAATCTCGGCCAGGCATGG 16055

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE

1 (bases 1 to 61412)

AUTHORS

Birren,B., Linton,L., Nusbaum,C. and Lander,E.

TITLE

Homo sapiens, clone RP11-104F3

JOURNAL

REFERENCE

2 (bases 1 to 61412)

AUTHORS

Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Beda,F.,
Boguslavskiy,L., Bouknight,B., Brown,A., Burkett,G., Castle,A.,
Chospel,Y., Colangelo,M., Collins,S., Collamore,A., Cooke,P.,
DeArellano,K., Dewar,K., Domino,M., Doyle,M., Fenestor,J.,
Ferreira,P., Fitzhugh,W., Forrest,C., Gage,D., Galagan,J.,
Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
Landers,T., Lehoczy,J., Levine,R., Lieu,C., Liu,G., Locke,K.,
Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,
McPheeters,R., Meldrim,J., Meneus,L., Morrow,J., Naylor,J.,
Norman,C.H., O'Connor,T., O'Donnell,P., Olivar,T.M., Peterson,K.,
Pierre,N., Pisani,C., Pollara,V., Raymond,C., Riley,R., Rothman,D.,
Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J.,
Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,
Zimmer,A. and Zody,M.

Direct Submission

TITLE

Submitted (16-JAN-2000) Whitehead Institute/MIT Center for Genome

JOURNAL

Research, 320 Charles Street, Cambridge, MA 02141, USA

COMMENT

On Jul 13, 2000 this sequence version replaced gi:6705607.

All repeats were identified using RepeatMasker:

Smith, A.F.A. & Green, P. (1996-1997)

<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L5374

Center clone name: 104_F_3

* NOTE: This record contains 63 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

* 1 896: contig of 896 bp in length
* 897 996: gap of 100 bp
* 997 1867: contig of 871 bp in length
* 1868 1967: gap of 100 bp
* 1968 2874: contig of 907 bp in length
* 2875 2974: gap of 100 bp
* 2975 3849: contig of 875 bp in length
* 3850 3949: gap of 100 bp
* 3950 4841: contig of 892 bp in length
* 4842 4941: gap of 100 bp
* 4942 5805: contig of 864 bp in length
* 5806 5905: gap of 100 bp
* 5906 6772: contig of 867 bp in length
* 6773 6872: gap of 100 bp
* 6873 7746: contig of 874 bp in length
* 7747 7846: gap of 100 bp
* 7847 8733: contig of 887 bp in length
* 8734 8833: gap of 100 bp
* 8834 9702: contig of 869 bp in length
* 9703 9802: gap of 100 bp
* 9803 10666: contig of 864 bp in length

* 10667 10766: gap of 100 bp
* 10767 11621: contig of 855 bp in length
* 11622 11721: gap of 100 bp
* 11722 12583: contig of 862 bp in length
* 12584 12683: gap of 100 bp
* 12684 13559: contig of 876 bp in length
* 13560 13659: gap of 100 bp
* 13660 14537: contig of 878 bp in length
* 14538 14637: gap of 100 bp
* 14638 15223: contig of 886 bp in length
* 15224 15624: gap of 100 bp
* 15624 16568: contig of 845 bp in length
* 16569 17437: contig of 869 bp in length
* 17438 17537: gap of 100 bp
* 17538 18397: contig of 860 bp in length
* 18398 18497: gap of 100 bp
* 18498 19355: contig of 858 bp in length
* 19356 19455: gap of 100 bp
* 19456 20366: contig of 911 bp in length
* 20367 20466: gap of 100 bp
* 20467 21361: contig of 895 bp in length
* 21362 21461: gap of 100 bp
* 21462 22361: contig of 900 bp in length
* 22362 22461: gap of 100 bp
* 22462 23361: contig of 900 bp in length
* 23362 23461: gap of 100 bp
* 23462 24327: contig of 866 bp in length
* 24328 24427: gap of 100 bp
* 24428 25319: contig of 892 bp in length
* 25320 25419: gap of 100 bp
* 25420 26258: contig of 839 bp in length
* 26259 26358: gap of 100 bp
* 26359 27250: contig of 892 bp in length
* 27251 27350: gap of 100 bp
* 27351 28229: contig of 879 bp in length
* 28230 28329: gap of 100 bp
* 28330 29174: contig of 845 bp in length
* 29175 29274: gap of 100 bp
* 29275 30135: contig of 861 bp in length
* 30136 30235: gap of 100 bp
* 30236 31099: contig of 864 bp in length
* 31100 31199: gap of 100 bp
* 31200 32108: contig of 909 bp in length
* 32109 32208: gap of 100 bp
* 32209 33061: contig of 853 bp in length
* 33062 34022: contig of 861 bp in length
* 34023 34122: gap of 100 bp
* 34123 34973: contig of 851 bp in length
* 34974 35073: gap of 100 bp
* 35074 35947: contig of 874 bp in length
* 35948 36047: gap of 100 bp
* 36048 36920: contig of 873 bp in length
* 36921 37020: gap of 100 bp
* 37021 37909: contig of 889 bp in length
* 37910 38009: gap of 100 bp
* 38010 38896: contig of 887 bp in length
* 38897 38996: gap of 100 bp
* 38997 39872: contig of 876 bp in length
* 39873 39972: gap of 100 bp
* 39973 40874: contig of 902 bp in length
* 40875 40974: gap of 100 bp
* 40975 41836: contig of 862 bp in length
* 41837 41936: gap of 100 bp
* 41937 42810: contig of 874 bp in length
* 42811 42910: gap of 100 bp
* 42911 43793: contig of 883 bp in length
* 43794 43893: gap of 100 bp
* 43894 44796: contig of 903 bp in length
* 44797 44896: gap of 100 bp
* 44897 45751: contig of 855 bp in length
* 45752 45851: gap of 100 bp

JOURNAL REFERENCE AUTHORS TITLE JOURNAL REFERENCE AUTHORS TITLE JOURNAL COMMENT

Unpublished 2 (bases 1 to 139280) Kaul,R.K., Olson,M.V., Raymond,C. and Haugen,E.D. Direct Submission Submitted (08-NOV-2001) Genome Center, University of Washington, Box 352145, Seattle, WA 98195, USA

3 (bases 1 to 139280) Kaul,R.K., Olson,M.V., Zhou,Y., James,R.A., Rouse,G., Wu,Z., Saenphimmachak,C., Phelps,K.A., Raymond,C. and Haugen,E.D. Direct Submission Submitted (27-APR-2002) Genome Center, University of Washington, Box 352145, Seattle, WA 98195, USA On Apr 27, 2002 this sequence version replaced gi:16799010.

Center: University of Washington Genome Center Center Code: UWGC Web site: http://www.genome.washington.edu Contact: uwgchgs@u.washington.edu Drafting Center: BCM ----- Project Information Center project name: chr-3 Center clone name: RP11-437M11 (bc0424) ----- Summary Statistics Sequencing vector: unknown; 33% of reads Chemistry: Dye-terminator ET; 76% of reads Assembly program: Phrap; version 0.990319 Consensus quality: 139214 bases at least Q40 Consensus quality: 139263 bases at least Q30 Insert size: 139280; sum-of-contigs Quality coverage: 13.0x in Q20 bases; sum-of-contigs -----

Overlapping Sequences: 5': RP11-219121 (UWGC:bc0309) AC092047, 130783-bp overlap 3': RP4-613B23 AC006059, 101219-bp clone overlap, of which 37845 bp is included in this submission -----

Sequence Quality Assessment: This entry has been annotated with sequence quality estimates computed by the Phrap assembly program. All manually edited bases have been reduced to quality zero. Quality levels above 40 are expected to have less than 1 error in 10,000 bp. Base-by-base quality values are not generally visible from the GenBank flat file format but are available as part of this entry's ASN.1 file.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest.

Sequence Validation: This sequence has been validated by Multiple Complete Digest fingerprinting. Comparison of the experimentally derived digest fragments with sequence-predicted fragments is given below. The electronically-digested sequence consists of both insert and vector, in order to accurately represent the entire circular BAC. Small fragments below a variable cutoff (approximately 400-800 bp) are not resolved in the fingerprint and hence do not appear in the table. There are no significant remaining discrepancies between the experimental and predicted values. Uniquely ordered fragments are separated by dashed lines.

BglII ----- EcoRI ----- HindIII ----- SeqDerMap FngPrnt SeqDerMap FngPrnt SeqDerMap FngPrnt

3334	3332	8696	8706	3211	3274	4501	4494	1539	1576	801	810
2067	2130	6	<800	6382	6509	1246	1196	9653	9683	2871	2870
7056	7041	1909	1855	512	<800	3506	3541	1139	1106	6483	6509
192	<800	9283	9141	449	<800	26	<800	4988	5046	475	<800
6704	6557	2412	2548	4573	4461	8	<800	6671	6733	13148	12835
1550	1550	1706	1657	3848	3827	2587	2707	10647	10644	2634	2667
592	<800	13289	13045	2253	2291	1786	1875	3923	3899	5084	5150
4161	4208	4390	4360	2063	2080	9945	9829	148	<800	3053	3135
646	<800	5556	5570	10655	10560	2384	2406	3510	3495	692	<800
11112	10883	59	<800	5319	5359	566	<800	5156	5046	11	<800
2333	2406	5316	5307	1692	1676	13468	14157	3082	3095	1217	1197
1312	1311	2688	2730	1757	1676	1736	1875	2190	2180	18955	19193
1020	1040	202	<800	2791	2870	1667	1699	3488	3495	-----	-----
9291	9095	1403	1383	1656	1676	4212	4208	16611	16955	-----	-----
4538	4494	5062	5046	7317	7406	429	<800	834	849	-----	-----
9281	9095	7618	7632	3977	4047	15	<800	1592	1657	-----	-----
2275	2406	3644	3628	356	<800	75	<800	1072	1053	-----	-----
4461	4494	1051	1053	10777	10560	1058	1040	11533	11404	-----	-----
621	<800	6005	6000	8263	8317	158	<800	1417	1504	-----	-----
1199	1196	102	<800	2997	3026	557	<800	-----	-----	-----	-----
4463	4494	6161	6223	1027	1025	Query Match 93.6%; Score 23.4; DB 8; Length 139280;					
9798	9829	3848	3899	9772	9719	Best Local Similarity 96.0%; Pred. No. 2.5;					
8361	8316	180	<800	460	<800	Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;					
327	<800	776	761	3648	3668	QY 1 AAAAAAAAAATCAGCGCCAGGCATGG 25					
1032	1040	1245	1197	16661	16928	Db 133475 AAAAAAAAAATCTCGGCCAGGCATGG 133451					
1918	1875	552	<800	5968	6078	RESULT 7					
1003	1040	2099	2180	150	<800	AC099405 142943 bp DNA linear PRI 11-JAN-2003					
7003	7041	3224	3219	6146	6078	AC099405 Papio anubis clone RP41-161K2, complete sequence.					
3271	3332	871	849	2639	2667	AC099405.3 GI:27597007					
331	<800	2537	2548	13119	12835	HTG.					
376	<800	745	<800	502	<800	Papio anubis (olive baboon)					
443	<800	4622	4657	110	<800	Papio anubis					
1790	1875	2156	2180	2363	2371	Eukaryota; Metazoa; Chordata; Craniata; Ayele, K., Beckstrom-Sternberg, S.M.,					
2604	2707	1211	1197	9088	8966	Bakariya, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S.,					
912	931	1686	1657	1847	1777	Pagitragan, C., Pearson, R., Portnoy, M.E., Prasad, A.,					
238	<800	9850	9683	1581	1525	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;					
						Cercopithecoidea; Cercopithecinae; Papio.					
						1 (bases 1 to 142943)					
						Akher, N., Anconellis, A., Ayele, K., Beckstrom-Sternberg, S.M.,					
						Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S.,					
						Carlag, K., Coleman, B., Engle, J., Granite, S., Guan, X., Gupta, J.,					
						Haghighi, P., Han, J., Hansen, N., Ho, S.-L., Idol, J.R., Karlins, E.,					
						Laric, P., Lee-Lin, S.-O., Legaspi, R., Maduro, O.L., Maduro, V.B.,					
						Marquies, E.H., Masello, C., Maskeri, B., McDownell, J.,					
						Pagitragan, C., Pearson, R., Portnoy, M.E., Prasad, A.,					
						Reddix-Dugue, N., Schandler, K., Schueler, M.G., Sison, C.,					
						Stantripop, S., Thomas, J.W., Thomas, P.J., Touchman, J.W., Vogt, J.L.,					
						Wetherby, K.D., Wiggins, L., Young, A. and Green, E.D.					

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homiidae; Pan.
 1 (bases 1 to 163473)
 Antonellis,A., Ayale,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
 Blakeley,R.W., Bouffard,G.G., Brinkley,C., Brooks,S., Cariaga,K.,
 Chu,G., Coleman,B., Coleman,H., Engle,J., Granite,S., Guan,X.,
 Gupta,J., Haghighi,P., Han,J., Hansen,N., Ho,S.-L., Hu,P.,
 Hurler,B., Idol,J.R., Karlins,E., Kwong,P., Laric,P., Lee-Lin,S.-O.,
 Legaapi,R., Maduro,Q.L., Maduro,V.B., Marquies,E.H., Mastello,C.,
 Maskeri,B., McDowell,J., Paguirigan,C., Pearson,R., Portnoy,M.E.,
 Prasad,A., Reddix-Dugue,N., Schandler,K., Schueler,M.G., Shah,K.,
 Sison,C., Stantripop,S., Thomas,J.W., Thomas,P.J., Tsipouri,V.,
 Vogt,J.L., Wetherby,K.D., Wiggins,L., Young,A. and Green,E.D.
 NISC Comparative Sequencing Initiative
 Unpublished
 2 (bases 1 to 163473)
 Green,E.D.
 Direct Submission
 Submitted (23-MAY-2003) NIH Intramural Sequencing Center, 8717
 Government Circle, Gaithersburg, MD 20877, USA
 3 (bases 1 to 163473)
 Green,E.D.
 Direct Submission
 Submitted (09-JUN-2003) NIH Intramural Sequencing Center, 8717
 Government Circle, Gaithersburg, MD 20877, USA
 On Jun 9, 2003 this sequence version replaced gi:31044283.
 ----- Genome Center
 Center: NIH Intramural Sequencing Center
 Center code: NISC
 Web site: <http://www.nisc.nih.gov>
 Contact: nisc.zoo@hgri.nih.gov
 ----- Project Information
 Center project name: esd
 Center clone name: 288H14

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics
 Sequencing vector: plasmid; n/a; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 162199 bases at least Q40
 Consensus quality: 162591 bases at least Q30
 Consensus quality: 162777 bases at least Q20
 Insert size: 149000; agarose-fp
 Insert size: 162873; sum-of-contigs
 Quality coverage: 10.65x in Q20 bases; agarose-fp
 Quality coverage: 9.74x in Q20 bases; sum-of-contigs

 * NOTE: This is a 'working draft' sequence. It currently consists of 7 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.
 1 49711: contig of 49711 bp in length
 49712 49811: gap of unknown length
 49812 93246: contig of 43435 bp in length
 93247 93347: gap of unknown length
 93347 126704: contig of 33358 bp in length

* 126705 126804: gap of unknown length
 * 126805 128476: contig of 1672 bp in length
 * 128477 128576: gap of unknown length
 * 128577 143981: contig of 15405 bp in length
 * 143982 144081: gap of unknown length
 * 144082 145826: contig of 1745 bp in length
 * 145827 145926: gap of unknown length
 * 145927 163473: contig of 17547 bp in length.
 FEATURES
 source
 1..163473
 /organism="Pan troglodytes"
 /mol_type="genomic DNA"
 /db_xref="taxon:9598"
 /clone="CH251-288H14"
 /clone_lib="CH251"
 /notes="clone overlaps with GenBank Accession Number AC096875 clone RP43-120N4 (center project name c12)"
 1..49711
 /notes="assembly_fragment
 clone_end:T7
 vector_side:left"
 49712..49811
 /estimated_length=unknown
 49812..93246
 /notes="assembly_fragment"
 93247..93346
 /estimated_length=unknown
 93347..126704
 /notes="assembly_fragment"
 103034..163473
 /notes="clone overlaps with GenBank Accession Number AC097009 clone RP43-169E8 (center project name cma)"
 126705..126804
 /estimated_length=unknown
 126805..128476
 /notes="assembly_fragment"
 128477..128576
 /estimated_length=unknown
 128577..143981
 /notes="assembly_fragment"
 143982..144081
 /estimated_length=unknown
 144082..145826
 /notes="assembly_fragment"
 145827..145926
 /estimated_length=unknown
 145927..163473
 /notes="assembly_fragment
 clone_end:SP6
 vector_side:right"
 ORIGIN
 Query Match 93.6%; Score 23.4; DB 14; Length 163473;
 Best Local Similarity 96.0%; Pred. No. 2.5;
 Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 AAAAAAAAAATCAGGCGCAGGCATGG 25
 |||||
 Db 45133 AAAAAAAAAATCAGGCGCAGGCATGG 45157
 RESULT 10
 AL135914/c
 LOCUS
 DEFINITION
 AL135914 165802 bp DNA linear PRI 18-MAY-2005
 Human DNA sequence from clone RP1-56L9 on chromosome 6 Contains a novel gene, the SOD2 gene for superoxide dismutase 2, mitochondrial, a pseudogene similar to heterogeneous nuclear ribonucleoprotein H, the gene for Wilms' tumor 1-associating protein (R1A0105) (WTAP), the ACAT2 gene for acetyl-Coenzyme A acetyltransferase 2 (acetoacetyl Coenzyme A thiolase), the TCPI gene for t-complex 1 (CTC1, Ccta, D6S230E), the 5' end of the MRPL18 gene for mitochondrial ribosomal protein L18, and 4 CpG


```
/evidence=not experimental
/rpt family="AluJb"
18465..16486
/evidence=not experimental
/rpt family="AT rich"
16492..16803
/evidence=not experimental
/rpt family="AluY"
16838..17049
/evidence=not experimental
/rpt family="AluJo"
17247..17332
/evidence=not experimental
/rpt family="MLT1K"
17851..17998
/evidence=not experimental
/rpt family="L2"
complement(18114..18296)
/evidence=not experimental
/rpt family="MIR"
complement(18503..18635)
/evidence=not experimental
/rpt family="L1M4"
complement(18851..18934)
/evidence=not experimental
/rpt family="L1P4"
19285..19356
/evidence=not experimental
/rpt family="L2"
19391..19697
/evidence=not experimental
/rpt family="AluSc"
19708..19987
/evidence=not experimental
/rpt family="AluSg"
19988..20022
/evidence=not experimental
/rpt family="(TAAA)n"
20035..20363
/evidence=not experimental
/rpt family="AluJo"
20558..20711
/evidence=not experimental
/rpt family="MLT1H"
20807..21027
/evidence=not experimental
/rpt family="AluJb"
21028..21062
/evidence=not experimental
/rpt family="AluY"
complement(21074..21375)
/evidence=not experimental
/rpt family="AluJo"
complement(21472..21611)
/evidence=not experimental
/rpt family="AluSg/x"
complement(21612..21623)
/evidence=not experimental
/rpt family="AluY"
21624..21643
/evidence=not experimental
/rpt family="(TTTA)n"
complement(21644..21913)
/evidence=not experimental
/rpt family="AluY"
21914..22185
/evidence=not experimental
/rpt family="AluJb"
22190..22505
/evidence=not experimental
/rpt family="AluSg"
22510..22812
/evidence=not experimental
```

```
repeat_region
/rpt family="AluJb"
22872..23141
/evidence=not experimental
/rpt family="AluSg1"
complement(23261..23563)
/evidence=not experimental
/rpt family="AluJo"
23741..24077
/evidence=not experimental
/rpt family="AluJb"
24095..24405
/evidence=not experimental
/rpt family="AluY"
24719..24774
/evidence=not experimental
/rpt family="(CATATA)n"
24915..25015
/evidence=not experimental
/rpt family="AT rich"
complement(25077..25132)
/evidence=not experimental
/rpt family="MLT1D"
25191..25245
/evidence=not experimental
/rpt family="MER66A"
complement(25556..25720)
/evidence=not experimental
/rpt family="L1M4"
25717..26058
/evidence=not experimental
/rpt family="L1ME1"
26150..26446
/evidence=not experimental
/rpt family="AluSx"
27486..27783
/evidence=not experimental
/rpt family="AluSx"
27982..28368
repeat_region
```

Query Match 93.6%; Score 23.4; DB 8; Length 168813;
Best Local Similarity 96.0%; Pred. No. 2.5;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAAAAAAAATCAGCGCGCATGG 25
|||||

Db 38340 AAAAAAAAAATCAGCGCGCATGG 38316
|||||

RESULT 12

AC146225/c

LOCUS

DEFINITION

AC146225

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

AC146225 171272 bp DNA linear PRI 18-DEC-2003
Pan troglodytes BAC clone RP43-32E8 from 7, complete sequence.

AC146225 2 GI:39725897
HTG.

Pan troglodytes (chimpanzee)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Pan.

1 (bases 1 to 171272)

Van Brunt, A., Bielicki, L. and Haglund, K.

The sequence of Pan troglodytes BAC clone RP43-32E8

Unpublished (2001)

2 (bases 1 to 171272)

Wilson, R.

Sequencing of Pan troglodytes

Unpublished (2001)

3 (bases 1 to 171272)

Wilson, R.K.

Direct Submission

Submitted (01-AUG-2003) Genetics, Genome Sequencing Center, 4444
Forest Park Parkway, St. Louis, MO 63108, USA

```

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
4 (bases 1 to 171272)
Wilson,R.K.
Direct Submission
Submitted (11-DEC-2003) Genetics, Genome Sequencing Center, 4444
Forest Park Parkway, St. Louis, MO 63108, USA
5 (bases 1 to 171272)
Wilson,R.
Direct Submission
Submitted (18-DEC-2003) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Dec 11, 2003 this sequence version replaced gi:3338178.
-----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu
Contact: submissions@wustl.edu
-----
Summary Statistics
Center project name: C_PT032E08
-----

NOTICE: This sequence may not represent the entire insert of this
clone. It may be shorter because we only sequence overlapping
clone sections once, or longer because we provide a small overlap
between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by
restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. Wes Warren,
Department of Genetics, Washington University, St. Louis MO. For
additional information about the map position of this sequence, see
http://genome.wustl.edu

SOURCE INFORMATION:
The RPCI-43 BAC Library has been constructed by Chung-Li Shu. DNA
was isolated from white blood cells obtained from a male chimpanzee
(Pan troglodytes, 'Cint', Verkes #C0471; birthdate: 6-6-80). The
clone and detailed information can be obtained from RegGen
(http://www.reggen.com) or Pieter de Jong and co-workers at
http://www.bacpac.chori.org.

NEIGHBORING SEQUENCE INFORMATION:
This sequence is the entire insert of the clone.
Location/Qualifiers
1..171272
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/chromosome="7"
/map="7"
/clone="RP43-32E8"
/clone_lib="RPCI-43"
98712..99115
/misc_feature
/note="Sequence derived from PCR product of project DNA."

ORIGIN
Query Match 93.6%; Score 23.4; DB 8; Length 171272;
Best Local Similarity 96.0%; Pred. NO. 2.5;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAAAAAAAATCAGCCCGCATGG 25
DB 5317 AAAAAAAAAATCAGCCCGCATGG 5293

RESULT 13
AC092041

```

```

LOCUS
DEFINITION
AC092041 Homo sapiens chromosome 3 clone RP11-132K4, complete sequence.
AC092041 AC018471
VERSION
AC092041.2 GI:16554366
KEYWORDS
HTG.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
REFERENCE
1 (bases 1 to 171461)
Kaul,R.K., Olson,M.V., Zhou,Y., James,R.A., Raymond,C. and
Haugen,E.D.
Direct Submission
Unpublished
2 (bases 1 to 171461)
Kaul,R.K., Olson,M.V., Raymond,C., Clendenning,J., Ivey,R.G. and
Haugen,E.D.
Direct Submission
Submitted (16-JUN-2001) Genome Center, University of Washington,
Box 352145, Seattle, WA 98195, USA
3 (bases 1 to 171461)
Kaul,R.K., Olson,M.V., Zhou,Y., James,R.A., Raymond,C. and
Haugen,E.D.
Direct Submission
Submitted (31-OCT-2001) Genome Center, University of Washington,
Box 352145, Seattle, WA 98195, USA
On Oct 31, 2001 this sequence version replaced gi:14475974.
-----
Center: University of Washington Genome Center
Center Code: UWGC
Web site: http://www.genome.washington.edu
Contact: uwgchgts@u.washington.edu
Drafting Center: BCM
-----
Project Information
Center project name: Chr-3
Center clone name: RP11-132K4 (bc0251)
-----
Summary Statistics
Sequencing vector: M13; L08821; 46% of reads
Sequencing chemistry: plasmid; L08752; 54% of reads
Chemistry: Dye-primer Bodipy; 21% of reads
Chemistry: Dye-terminator Eri; 41% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 170617 bases at least Q40
Consensus quality: 171425 bases at least Q30
Consensus quality: 171461 bases at least Q20
Insert size: 172100; 9.1% error; agarose-fp
Insert size: 169288; sum-of-contigs
Quality coverage: 6.4x in Q20 bases; agarose-fp
Quality coverage: 6.5x in Q20 bases; sum-of-contigs
-----
Overlapping Sequences:
5': RP11-296G21 (UWGC:bc0347) AC022064
3': Mapping in progress
-----
Sequence Quality Assessment:
This entry has been annotated with sequence quality
estimates computed by the Phrap assembly program.
All manually edited bases have been reduced to quality zero.
Quality levels above 40 are expected to have less than
1 error in 10,000 bp.
Base-by-base quality values are not generally visible from the
GenBank flat file format but are available as part
of this entry's ASN.1 file.

This sequence was finished as follows unless otherwise noted:
all regions were either double-stranded or sequenced with an
alternate chemistry or covered by high quality data (i.e., Phred
quality >= 30); an attempt was made to resolve all sequencing
problems, such as compressions and repeats; all regions were
covered by at least one plasmid subclone or more than one M13
subclone; and the assembly was confirmed by restriction digest.

```

Sequence Validation:

This sequence has been validated by Multiple Complete Digest fingerprinting. Comparison of the experimentally derived digest fragments with sequence-predicted fragments is given below. The electronically-digested sequence consists of both insert and vector, in order to accurately represent the entire circular BAC. Small fragments below a variable cutoff (approximately 400-800 bp) are not resolved in the fingerprint and hence do not appear in the table. There are no significant remaining discrepancies between the experimental and predicted values. Uniquely ordered fragments are separated by dashed lines.

EcoRI				HindIII				BglII			
SeqDerMap	FngrPrnt	SeqDerMap	FngrPrnt	SeqDerMap	FngrPrnt	SeqDerMap	FngrPrnt	SeqDerMap	FngrPrnt	SeqDerMap	FngrPrnt
8696	8718	6678	6413	4913	4847						
6	<800	6382	6413	2067	2044						
2156	2188	512	<800	6965	7043						
1211	1199	449	<800	331	<800						
1686	1660	3307	3321	376	<800						
9850	10016	5968	5988	443	<800						
1539	1512	150	<800	1790	1807						
9653	10016	6146	6413	2604	2723						
1139	1096	2639	2658	912	932						
4989	5165	13120	13188	238	<800						
6671	7108	502	<800	4501	4488						
10647	10016	110	<800	1246	1203						
3923	3833	2363	2357	3506	3476						
148	<800	9088	9105	26	<800						
3510	3471	1847	1776	8	<800						
5156	5165	1581	1535	2587	2618						
3082	3060	801	822	1786	1807						
2190	2188	2871	2888	9946	10100						
3488	3471	6483	6413	2384	2295						
16610	16826	475	<800	566	<800						
834	851	13148	13188	13468	13603						
1592	1584	2634	2658	1736	1807						
1072	1096	5083	5080	1667	1625						
11533	11673	3053	3133	4212	4175						
1414	1403	692	<800	429	<800						
13506	13186	11	<800	15	<800						
13438	13186	1217	1225	75	<800						
7021	7108	18955	19298	1058	1074						

4626	4653	7378	7543	158	<800
373	<800	2204	2219	557	<800
660	<800	3558	3549	1941	1945
2077	2188	10428	10124	1851	1807
192	<800	6333	6413	1358	1325
565	<800	2867	2888	1100	1074
2741	2713	7646	7543	10141	10100
2032	2053	2371	2357	2785	2974
14931	15070	1276	1225	930	932
5203	5165	1104	1106	1836	1807
		15297	14858	2679	2832
		3433	3549	2950	2974
				3238	3265
				4244	4175
				106	<800
				5422	5449
				3226	3265
				3810	3851
				2893	2974
				1623	1625
				14167	13603
				8682	8622
				1609	1625
				12735	13603
				12350	11745
				2422	2478
				5492	5449

Query Match 93.6%; Score 23.4; DB 8; Length 171461;
 Best Local Similarity 96.0%; Pred. No. 2.5;
 Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAAAAAAAATCAGGCCAGGCATGG 25
 |||||
 Db 136541 AAAAAAAAAATCAGGCCAGGCATGG 136565

RESULT 14

AC010653
 LOCUS AC010653 184515 bp DNA linear PRI 06-AUG-2002
 DEFINITION Homo sapiens chromosome 16 clone RP11-417N10, complete sequence.
 ACCESSION AC010653
 VERSION AC010653.10 GI:22122867
 KEYWORDS HTG.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.

REFERENCE
TITLES
DOE Joint Genome Institute.
Sequencing of Human Chromosome 16

JOURNAL
Unpublished

REFERENCE
AUTHORS
DOE Joint Genome Institute.

JOURNAL
Direct Submission

TITLE
Submitted (16-SEP-1999) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA

REFERENCE
AUTHORS
DOE Joint Genome Institute.

JOURNAL
Direct Submission

TITLE
Submitted (06-AUG-2002) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA

COMMENT
On Aug 6, 2002 this sequence version replaced gi:17737037.

Sequence Quality Assessment:

This entry has been annotated with sequence quality

estimates computed by the Phrap assembly program.

All manually edited bases have been reduced to quality zero.

Quality levels above 40 are expected to have less than

1 error in 10,000 bp.

Base-by-base quality values are not generally visible from the

GenBank flat file format but are available as part

of this entry's ASN.1 file.

Sequence Quality Assessment:

This entry has been annotated with sequence quality

estimates computed by the Phrap assembly program.

All manually edited bases have been reduced to quality zero.

Quality levels above 40 are expected to have less than

1 error in 10,000 bp.

Base-by-base quality values are not generally visible from the

GenBank flat file format but are available as part

of this entry's ASN.1 file.

FEATURES

source

1. 184515

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="16"

/clone="RP11-417N10"

ORIGIN

Query Match 93.6%; Score 23.4; DB 8; Length 184515;

Best Local Similarity 96.0%; Pred. No. 2.5;

Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AAAAAAAAAATCAGGCCAGGCATGG 25

|||||

Db 176048 AAAAAAAAAATCAGGCCAGGCATGG 176072

|||||

AC006059

AC006059

AC006059.3 GI:4544348

HTG.

AC006059

HTG.

AC006059

HTG.

AC006059

HTG.

AC006059

HTG.

AC006059

HTG.

He, X., Hernandez, J., Jackson, L.E., Kondejewski, N., Leal, B., Lichtarge, O., Liu, W., Logan, O., Lu, J., Martinez, C., Moore, S., Moorhage, T., Nguyen, N., Oswal, G., Pampall, L.R., Parish, B.J., Perez, L.M., Rashid, N.D., Rives, C.M., Scherer, S.E., Shen, H., Simon, M.L., Vo, Q.K., Wei, Y., Williamson, A.L., Worley, K., Zhou, X., Naylor, S.L. and Gibbs, R.A.

Direct Submission

Unpublished

2 (bases 1 to 185161)

Worley, K.C.

Direct Submission

Submitted (26-NOV-1998) Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 185161)

Worley, K.C.

Direct Submission

Submitted (30-MAR-1999) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

4 (bases 1 to 185161)

Worley, K.C.

Direct Submission

Submitted (01-APR-1999) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On Mar 30, 1999 this sequence version replaced gi:4263627.

INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email

gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.

ANNOTATION OF FEATURES:

STSs are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences.

Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL:

<http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

QUALSTAT-REPORT-----

----- Summary Statistics -----
Contig length: 185161
Phrap values in estimate: 185162
Average error rate (BCM-Phrap estimate): 0.000104921
Fraction of Phrap values less than 40: 0.0374023
Number of consensus changing edits: 20
Number of N's in consensus: 0

Consensus changing edits		
Position	Original+Context	Edited+Context
28864	actcaggcag(t)tcgagagcca	actcaggcag(g)gcgagagcca
28865	ctcaggcagt(t)cgagagccag	ctcaggcagg(g)cgagagccag
32674	ctcaggcgag(n)aacagcctag	ctcaggcgag(c)aacagcctag
53006	cacggtgcca(a)gcttattttt	cacggtgcca(g)gcttattttt
84131	ctccccaac(n)ctcccaccc	ctccccaac(c)ctcccaccc
100219	ctggaacct(n)cacctccttt	ctggaacct(g)cacctccttt
109068	tccctcctct(n)ccccctctcc	tccctcctct(c)ccccctctcc
109069	ccccctctn(n)ccccctctcc	ccccctctc(t)ccccctctcc
149456	ctttctctc(n)ggaaaaaaa	ctttctctc(t)ggaaaaaaa
149636	aataaataat(n)caaacagcca	aataaataat(c)caaacagcca
158118	ttcaacaagta(n)aaaagagcat	ttcaacaagta(a)aaaagagcat
163739	tgcacctcag(n)attccagcct	tgcacctcag(c)attccagcct
167565	gagggcgggg(n)tgagacgaag	gagggcgggg(c)tgagacgaag
169144	tttgtttcca(n)tnatctgtgt	tttgtttcca(a)tnatctgtgt
169146	ttgttccant(n)atctgtgtct	ttgttccaa(g)atctgtgtct
169213	acattacatg(n)ncggcgctgg	acattacatg(g)ncggcgctgg
169214	cattacatgn(n)cgggcgctgg	cattacatg(g)cgggcgctgg
176105	ctacaataat(n)aaaattagca	ctacaataat(a)aaaattagca
178686	gaattggttt(n)gtggaagaca	gaattggttt(c)gtggaagaca
183324	gttgttttgt(n)ttgttttgtt	gttgttttgt(t)ttgttttgtt

Distribution of Quality < 40 Bases

# bases	Phrap Value Range									
	5	10	15	20	25	30	35	40		
1000							*	*		
900							*	*		
800							*	*		
700							*	*		
600							*	*		
500							*	*		
400						*	*	*		
300					*	*	*	*		
200				*	*	*	*	*		
100	*	*	*	*	*	*	*	*		
0							*	*		

Version: 1.01 qxf0.

FEATURES	source	Location/Qualifiers	
		1.	.185161
repeat_region	linear	/organism="Homo sapiens"	
		/mol_type="genomic DNA"	
		/db_xref="taxon:9606"	
		/clone="RPC14-613B23"	
		complement(1..79)	
		/rpt_family="AluS"	
		complement(80..484)	
		/rpt_family="L1M4"	
		complement(485..785)	
		/rpt_family="AluS"	
		complement(786..1830)	
		/rpt_family="L1M4"	
		complement(1851..1932)	
		/rpt_family="AluJ/FRAM"	
repeat_region	linear	2846..2885	
		/rpt_family="MER5B"	
		2886..3070	
		/rpt_family="AluSp"	
		complement(3263..3380)	
		/rpt_family="MIR"	
		4087..4629	
		/rpt_family="MER4B"	
		4669..4970	
		/rpt_family="AluSx"	
		5120..5508	

repeat_region	/rpt_family="MER57B"	complement(5509..5687)
repeat_region	/rpt_family="AluJb"	5712..6160
repeat_region	/rpt_family="MER77"	6295..6461
repeat_region	/rpt_family="MER5B"	complement(6858..7224)
repeat_region	/rpt_family="L2"	7524..7663
repeat_region	/rpt_family="L1ME1"	complement(7666..7699)
repeat_region	/rpt_family="AT rich"	complement(7700..7790)
repeat_region	/rpt_family="AluSg/x"	complement(7791..7832)
repeat_region	/rpt_family="AluSc"	complement(7831..7876)
repeat_region	/rpt_family="AluSg/x"	complement(7877..8139)
repeat_region	/rpt_family="AluSx"	complement(8140..8161)
repeat_region	/rpt_family="(TAAA)n"	complement(8162..8463)
repeat_region	/rpt_family="AluSg"	8464..9146
repeat_region	/rpt_family="L1ME1"	complement(9148..9436)
repeat_region	/rpt_family="AluJb"	9487..9712
repeat_region	/rpt_family="L1ME3"	complement(9715..9841)
repeat_region	/rpt_family="MER5B"	complement(10378..10411)
repeat_region	/rpt_family="(CA)n"	10424..10497
repeat_region	/rpt_family="L1M4"	complement(10482..10751)
repeat_region	/rpt_family="L1M4"	

Query Match 93.6%; Score 23.4; DB 8; Length 185161;
Best Local Similarity 96.0%; Pred. No. 2.5;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy	1	AAAAAAAAATCAGCGCCAGCATGG	25
Db	153083	AAAAAAAAATCTCGCCAGCATGG	153107

RESULT 16	
AC096875	
LOCUS	AC096875.2 GI:21070677
DEFINITION	HTG; HTGS PHASE2; HTGS_DRAFT.
ACCESSION	AC096875
VERSION	HTG; HTGS PHASE2; HTGS_DRAFT.
KEYWORDS	Pan troglodytes (Chimpanzee)
SOURCE	Pan troglodytes
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Pan.
REFERENCE	1 (bases 1 to 187530)
AUTHORS	Akhter,N., Antonellis,A., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B., Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S., Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P., Hansen,N., Ho,S.-L., Idol,J.R., Karlins,E., Laric,P., Lee-Lin,S.-Q., Legaspi,R., Maduro,O.L., Maduro,V.B., Margulies,E.H., Mastello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C., McDowell,J., Faguirigan,C., Pearson,R., Portnoy,M.E., Prasad,A., Schueler,M.G., Stantropop,S., Thomas,J.W., Thomas,P.J., Touchman,J.W., Tsurgon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L., Young,A., Zhang,L.-H. and Green,E.D.

TITLE NISC Comparative Sequencing Initiative
 JOURNAL Unpublished
 REFERENCE 2. (bases 1 to 187530)
 AUTHORS Green, E.D.
 TITLE Direct Submission
 JOURNAL Submitted (29-SEP-2001) NIH Intramural Sequencing Center, 8717
 Grovemont Circle, Gaithersburg, MD 20877, USA
 REFERENCE 3. (bases 1 to 187530)
 AUTHORS Green, E.D.
 TITLE Direct Submission
 JOURNAL Submitted (22-MAY-2002) NIH Intramural Sequencing Center, 8717
 Grovemont Circle, Gaithersburg, MD 20877, USA
 COMMENT On May 22, 2002 this sequence version replaced gi:15808565.
 ----- Genome Center
 Center: NIH Intramural Sequencing Center
 Center code: NISC
 Web site: <http://www.nisc.nih.gov>
 Contact: nisc.zoo@nhgri.nih.gov
 ----- Project Information

 Center project name: clz
 Center clone name: 120N04

This sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics

Sequencing vector: plasmid; n/a; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 186653 bases at least Q40
 Consensus quality: 186859 bases at least Q30
 Consensus quality: 187025 bases at least Q20
 Insert size: 154000; agarose-fp
 Insert size: 187130; sum-of-contigs
 Quality coverage: 11.97x in Q20 bases; agarose-fp
 Quality coverage: 9.85x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently consists of 5 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.

* This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.

* 1 5374: contig of 5374 bp in length
 * 5375 5474: gap of unknown length
 * 5475 57679: contig of 5205 bp in length
 * 57680 57779: gap of unknown length
 * 57780 99879: contig of 42100 bp in length
 * 99880 99979: gap of unknown length
 * 99980 121286: contig of 21307 bp in length
 * 121287 121386: gap of unknown length
 * 121387 187530: contig of 66144 bp in length.

FEATURES

source
 1. 187530
 /organism="Pan troglodytes"
 /mol_type="Genomic DNA"
 /db_xref="taxon:9598"
 /clone="RP43-120N4"
 /clone_lib="RP43"
 misc_feature
 1. 5374
 /note="assembly_fragment"

clone_end:T7
 vector_side:left
 5375..5474
 /estimated_length=unknown
 misc_feature
 5475..57679
 /note="assembly_fragment"
 gap
 57680..57779
 /estimated_length=unknown
 misc_feature
 57780..99879
 /note="assembly_fragment"
 gap
 99880..99979
 /estimated_length=unknown
 misc_feature
 99980..121286
 /note="assembly_fragment"
 gap
 121287..121386
 /estimated_length=unknown
 misc_feature
 121387..187530
 /note="assembly_fragment"
 clone_end:SP6
 vector_side:right

ORIGIN
 Query Match 93.6%; Score 23.4; DB 14; Length 187530;
 Best Local Similarity 96.0%; Pred. NO. 2.5;
 Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAAAAAAAATCAGGCGCAGGCATGG 25
 |||||
 Db 171950 AAAAAAAAAATCAGGCGCAGGCATGG 171974

RESULT 17
 AC009097/c 199471 bp DNA linear PRI 19-MAR-2003
 LOCUS Homo sapiens chromosome 16 clone RP11-43215, complete sequence.
 DEFINITION AC009097
 ACCESSION AC009097
 VERSION AC009097.9 GI:29124044
 KEYWORDS HTG.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Homo.
 REFERENCE 1. (bases 1 to 199471)
 AUTHORS DOE Joint Genome Institute, Stanford Human Genome Center and Los
 Alamos National Laboratory.
 TITLE Direct Submission
 JOURNAL Unpublished
 REFERENCE 2. (bases 1 to 199471)
 AUTHORS DOE Joint Genome Institute.
 TITLE Direct Submission
 JOURNAL Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint
 Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
 REFERENCE 3. (bases 1 to 199471)
 AUTHORS DOE Joint Genome Institute.
 TITLE Direct Submission
 JOURNAL Submitted (27-MAR-2002) Production Sequencing Facility, DOE Joint
 Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
 REFERENCE 4. (bases 1 to 199471)
 AUTHORS DOE Joint Genome Institute.
 TITLE Direct Submission
 JOURNAL Submitted (29-MAR-2002) Production Sequencing Facility, DOE Joint
 Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
 REFERENCE 5. (bases 1 to 199471)
 AUTHORS DOE Joint Genome Institute.
 TITLE Direct Submission
 JOURNAL Submitted (01-MAY-2002) Production Sequencing Facility, DOE Joint
 Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
 REFERENCE 6. (bases 1 to 199471)
 AUTHORS DOE Joint Genome Institute, Stanford Human Genome Center and Los
 Alamos National Laboratory.
 TITLE Direct Submission
 JOURNAL Submitted (19-MAR-2003) DOE Joint Genome Institute, 2800 Mitchell

Drive, Walnut Creek, CA 94598, USA
On Mar 19, 2003 this sequence version replaced gi.20376957.
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Laboratory
National Laboratory
www-shgc.stanford.edu
Quality: Phrap Quality >=40 100% of Sequence;
Estimated Total Number of Errors is 0.
Location/Qualifiers
1. .199471
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="16"
/clone="RP11-43215"

ORIGIN
Query Match 93.6%; Score 23.4; DB 8; Length 199471;
Best Local Similarity 96.0%; Pred. No. 2.5;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAAAAAAAATCAGCGCCAGGCATGG 25
|||||||
Db 147272 AAAAAAAAAATCAGCGCCAGGCATGG 147248
|||||||

RESULT 18
AL590137 328439 bp DNA linear HTG 29-JUN-2002
LOCUS Homo sapiens chromosome 1 clone RP11-23017, WORKING DRAFT SEQUENCE,
DEFINITION 62 unordered pieces.
ACCESSION AL590137 AC026065
VERSION AL590137.7 GI:21665915
KEYWORDS HTG; HTGS PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
1 (bases 1 to 328439)
McLay, K.
Direct Submission
Submitted (28-JUN-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Jul 2, 2002 this sequence version replaced gi:21530915.
Draft Sequence Produced by Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
http://genome.wustl.edu/gsc/index.shtml
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: BA23017
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: M13; M7815; 66% of reads
Sequencing vector: plasmid; L08752; 33% of reads
Chemistry: Dye-terminator; 5% of reads
Chemistry: Dye-primer-amersham; 64% of reads
Chemistry: Dye-terminator ET-amersham; 0% of reads Chemistry:
Dye-terminator Big Dye; 29% of reads
Consensus quality: 299728 bases at least Q40
Consensus quality: 310124 bases at least Q30
Consensus quality: 316292 bases at least Q20
Insert size: 322339; sum-of-contigs
Insert size: 163622; 9.0% error; agarose-fp
Quality coverage: 3.54x in Q20 bases; sum-of-contigs Quality
Coverage: 7.39x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently
* consists of 62 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1
* 11330: contig of 11330 bp in length
* 11331 11330: gap of 100 bp
* 11431 13516: contig of 2086 bp in length
* 13517 13516: gap of 100 bp
* 13617 16748: contig of 3132 bp in length
* 16749 16748: gap of 100 bp
* 16848 22447: contig of 5599 bp in length
* 22448 22547: gap of 100 bp
* 22548 25403: contig of 2856 bp in length
* 25404 25503: gap of 100 bp
* 25504 27672: contig of 2169 bp in length
* 27673 27772: gap of 100 bp
* 27773 31223: contig of 3451 bp in length
* 31224 31323: gap of 100 bp
* 31324 39821: contig of 8498 bp in length
* 39822 39921: gap of 100 bp
* 39922 42084: contig of 2163 bp in length
* 42085 42184: gap of 100 bp
* 42185 52164: contig of 9980 bp in length
* 52165 52264: gap of 100 bp
* 52265 55940: contig of 3676 bp in length
* 55941 58147: contig of 2107 bp in length
* 58148 58247: gap of 100 bp
* 58248 61880: contig of 3633 bp in length
* 61881 64762: contig of 2782 bp in length
* 64763 64862: gap of 100 bp
* 64863 67436: contig of 2574 bp in length
* 67437 67536: gap of 100 bp
* 67537 70329: contig of 2793 bp in length
* 70330 70429: gap of 100 bp
* 70430 74265: contig of 3836 bp in length
* 74266 74365: gap of 100 bp
* 74366 78259: contig of 3894 bp in length
* 78260 78359: gap of 100 bp
* 78360 89197: contig of 10838 bp in length
* 89198 89297: gap of 100 bp
* 89298 92173: contig of 2876 bp in length
* 92174 92273: gap of 100 bp
* 92274 94940: contig of 2667 bp in length
* 94941 95040: gap of 100 bp
* 95041 99683: contig of 4643 bp in length
* 99684 99783: gap of 100 bp
* 99784 103630: contig of 3847 bp in length
* 103631 103730: gap of 100 bp
* 103731 108942: contig of 5212 bp in length
* 108943 109042: gap of 100 bp
* 109043 112029: contig of 2987 bp in length
* 112030 112129: gap of 100 bp
* 112130 116142: contig of 4013 bp in length
* 116143 116242: gap of 100 bp
* 116243 123932: contig of 7590 bp in length
* 123933 123932: gap of 100 bp
* 123934 127609: contig of 3677 bp in length
* 127610 127709: gap of 100 bp
* 127710 135211: contig of 7502 bp in length
* 135212 135311: gap of 100 bp
* 135312 137629: contig of 2318 bp in length
* 137630 137729: gap of 100 bp
* 137730 145291: contig of 7562 bp in length
* 145292 145391: gap of 100 bp
* 145392 149032: contig of 3641 bp in length
* 149033 149132: gap of 100 bp

* 149133 155039: contig of 5907 bp in length
* 155040 155139: gap of 100 bp
* 155140 158523: contig of 3384 bp in length
* 158524 158623: gap of 100 bp
* 158624 161127: contig of 2504 bp in length
* 161128 161227: gap of 100 bp
* 161228 165679: contig of 4452 bp in length
* 165680 165779: gap of 100 bp
* 165780 171609: contig of 5830 bp in length
* 171610 171709: gap of 100 bp
* 171710 174717: contig of 3008 bp in length
* 174718 174817: gap of 100 bp
* 174818 202623: contig of 27806 bp in length
* 202624 202723: gap of 100 bp
* 202724 207580: contig of 4857 bp in length
* 207581 207680: gap of 100 bp
* 207681 219602: contig of 11922 bp in length
* 219603 219702: gap of 100 bp
* 219703 222053: contig of 2351 bp in length
* 222054 222153: gap of 100 bp
* 222154 227017: contig of 4864 bp in length
* 227018 227117: gap of 100 bp
* 227118 232205: contig of 5088 bp in length
* 232206 232305: gap of 100 bp
* 232306 235573: contig of 3268 bp in length
* 235574 235673: gap of 100 bp
* 235674 237804: contig of 2131 bp in length
* 237805 237904: gap of 100 bp
* 237905 240134: contig of 2230 bp in length
* 240135 252023: contig of 11789 bp in length
* 252024 252123: gap of 100 bp
* 252124 252491: contig of 7368 bp in length
* 252492 259492: gap of 100 bp
* 259493 259592: contig of 2418 bp in length
* 262010 262109: gap of 100 bp
* 262110 264611: contig of 2502 bp in length
* 264612 264711: gap of 100 bp
* 264712 289527: contig of 4816 bp in length
* 289528 289628: gap of 100 bp
* 275447 275547: contig of 5820 bp in length
* 275548 282359: contig of 6812 bp in length
* 282360 282458: gap of 100 bp
* 282459 286378: contig of 3919 bp in length
* 286379 286478: gap of 100 bp
* 286479 291251: contig of 4773 bp in length
* 291252 291351: gap of 100 bp
* 291352 298416: contig of 7065 bp in length
* 298417 298516: gap of 100 bp
* 298517 305839: contig of 7323 bp in length
* 305840 305939: gap of 100 bp
* 305940 309254: contig of 3315 bp in length
* 309255 309354: gap of 100 bp
* 309355 312006: contig of 2652 bp in length
* 312007 312106: gap of 100 bp
* 312107 325800: contig of 13694 bp in length
* 325801 325900: gap of 100 bp
* 325901 328439: contig of 2539 bp in length.
FEATURES
 source
 1..328439
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="1"
 /clone="RP11-23017"
 /clone.lib="RPC1-11.1"
 1..11330
 /note="assembly fragment:01560
 fragment_chain:1"
 11431..13516
 /note="assembly fragment:02337
 fragment_chain:1"

misc_feature 13617..16748
 /note="assembly fragment:00007
 fragment_chain:1"
 16849..22447
 /note="assembly fragment:03052
 fragment_chain:2"
 22548..25403
 /note="assembly fragment:00504
 fragment_chain:2"
 25504..27672
Query Match 93.6%; Score 23.4; DB 14; Length 328439;
Best Local Similarity 96.0%; Pred. No. 2.5;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 AAAAAAAAAATCAGGCCAGGCATGG 25
 |||||
Db 13205 AAAAAAAAAATCAGGCCAGGCATGG 13229
RESULT 19
AB030316/c
LOCUS AB030316 12593 bp DNA linear ROD 19-OCT-2004
DEFINITION Mus musculus Pig-n gene for phosphatidylinositolglycan class N,
complete cds.
ACCESSION AB030316 AB030309 AB030310 AB030311 AB030312 AB030313
AB030314 AB030315
VERSION AB030316.2 GI:60391425
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Hong, Y., Maeda, Y., Watanabe, R., Ohishi, K., Mishkind, M., Riezman, H.
and Kinoshita, T.
TITLE pig-n, a mammalian homologue of yeast Mcd4p, is involved in
transferring phosphoethanolamine to the first mannose of the
glycosylphosphatidylinositol
J. Biol. Chem. 274 (49), 35099-35106 (1999)
JOURNAL PUBLISHED 10574991
REFERENCE 2 (bases 1 to 12593)
AUTHORS Hong, Y.
TITLE Direct Submission
JOURNAL Submitted (23-JUL-1999) Yeongjin Hong, Research Institute for
Microbial Diseases, Department of Immunoregulation; 3-1,
Yamada-Oka, Suita, Osaka 565-0871, Japan
(E-mail:yjhong@biken.osaka-u.ac.jp, Tel:81-6-879-8329,
Fax:81-6-875-5233)
COMMENT On or before Mar 1, 2005 this sequence version replaced gi:5668561,
gi:5668562, gi:5668563, gi:5668564, gi:5668565, gi:5668566,
gi:5668567, gi:5668568, gi:5668569.
FEATURES
 source
 1..12593
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /tissue_type="liver"
 /tissue_lib="mouse 129 liver genomic library"
 237..11714
 /gene="Pig-n"
 join(237..399,1212..1285,1572..1691,2728..2976,3792..3913,
 4547..4645,5355..5461,6225..6349,7499,7629,8168..8284,
 8668..8708,9049..9108,10944..11036,11655..11714)
 /gene="Pig-n"
 237..399
 /gene="Pig-n"
 /product="phosphatidylinositolglycan class N"
 /number=1
 945..1044
 /estimated_length=unknown
 1212..1285
gene
mRNA
exon
gap
exon

/gene="Pig-n"
/product="phosphatidylinositolglycan class N"
/number=2
exon 1572. .1691
/gene="Pig-n"
/number=3
exon 2728. .2976
/gene="Pig-n"
/number=4
CDS join(2756. .2976,3792. .3913,4547. .4645,5355. .5461,
6225. .6349,7499. .7629,8168. .8284,8668. .8708,9049. .9108,
10944. .11036,11655. .11714)
/gene="Pig-n"
/note="homology to yeast Saccharomyces cerevisiae MCD4"
/codon_start=1
/product="phosphatidylinositolglycan class N"
/protein_id="BAA82663.1"
/db_xref="GI:5668571"
translation="MLLPALGLLIHFVFFASIFDIYFTSLVHGMTPQFTPLPPAK
RLVLFVADGLRADTLIELDEGNSRAPFIRNVMHSGVSHTRVPTESRPGHVALI
AGFYEDVSAVAGKVENPEFDSLFNSKFTWSGSPDILPMFAGAGSDHVTIYSYD
AQREFGADATKLTDTVDKVDKDFDFAARNNSQSLFTKVNEEKVVFLLHLLGIDTNGH
AHRPSRREYKIKKVPQNSAQOYDDEFLKRWLENKRWKRDVNOADIAPLMASLIGV
STLTPTFTWGAGIKFPQNSAQOYDDEFLKRWLENKRWKRDVNOADIAPLMASLIGV
PPLNSVGILPVGVLNNTGLFKAESMFTNAVQLLEQFKVQTKKEATLPLFLTFFK"
3792. .3913
exon /gene="Pig-n"
/number=5
exon 4547. .4645
/gene="Pig-n"
/number=6
gap 5081. .5180
/estimated_length=unknown
exon 5355. .5461
/gene="Pig-n"
/product="phosphatidylinositolglycan class N"
/number=7
gap 5927. .6026
/estimated_length=unknown
exon 6225. .6349
/gene="Pig-n"
/product="phosphatidylinositolglycan class N"
/number=8
gap 6893. .6992
/estimated_length=unknown
exon 7499. .7629
/gene="Pig-n"
/product="phosphatidylinositolglycan class N"
/number=9
gap 7873. .7972
/estimated_length=unknown
exon 8168. .8284
/gene="Pig-n"
/product="phosphatidylinositolglycan class N"
/number=10
gap 8373. .8472
/estimated_length=unknown
exon 8668. .8708
/gene="Pig-n"
/product="phosphatidylinositolglycan class N"
/number=11
gap 8931. .9030
/estimated_length=unknown
exon 9049. .9108
/gene="Pig-n"
/product="phosphatidylinositolglycan class N"
/number=12
gap 9781. .9880
/estimated_length=unknown
exon 10944. .11036
/gene="Pig-n"
/number=13
exon 11655. .11710

/gene="Pig-n"
/number=14
ORIGIN
Query Match 89.6%; Score 22.4; DB 9; Length 12593;
Best Local Similarity 92.0%; Pred. No. 7.6;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 AAAAAAAAAATCACGCCAGGCATGG 25
|||||||
Db 3290 AAAAAAAAAATRANGCCAGGCATGG 3266
|||||||
RESULT 20
BV092806/c
LOCUS BV092806 394 bp DNA linear STS 15-OCT-2003
DEFINITION RPAMMSEQ0004730 Roche Palo Alto Mus musculus STS genomic, sequence
tagged site.
ACCESSION BV092806
VERSION BV092806.1 GI:37670285
KEYWORDS STS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 394)
AUTHORS Usuka,J., Liao,G., Cheng,J., Nguyen,A., Bach,C., Puech,A.,
McPherson,J.D., Foernzler,D. and Peltz,G.
TITLE Mus musculus SNPs
JOURNAL Unpublished (2003)
COMMENT
Contact: Jonathan Usuka
Roche Palo Alto Genetics and Genomics Department
Roche Palo Alto
3431 Hillview Ave, Mailstop S3-1, Palo Alto, CA 94024, USA
Tel: 6508555807
Email: Jonathan.Usuka@roche.com
Primer A: No primer submitted.
Primer B: No primer submitted.
Location/Qualifiers
FEATURES
source
1. .394
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/map="1-13610-13406-CAA01163581.1.1.19206"
/clone_lib="Roche Palo Alto"
/note="SNPs developed from assay sequences derived from 15
different strains
of mice (as of October 1, 2003). Those strains include
A/J, A/HeJ,
129/Sv, AKR/J, B10.D2-H2/OsNj, BALB/cByJ, BALB/cJ,
C3H/HeJ, C57BL/6J,
CAST/Ei, DBA/2J, MRL/MpJ, NZB/BinJ, NZW/Lac, SPRET/Ei.
"
ORIGIN
Query Match 87.2%; Score 21.8; DB 10; Length 394;
Best Local Similarity 92.0%; Pred. No. 15;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 AAAAAAAAAATCACGCCAGGCATGG 25
|||||||
Db 163 AAAAAAAAAATCTTGGCCAGGCATGG 139
|||||||
RESULT 21
BV099496/c
LOCUS BV099496 398 bp DNA linear STS 14-JAN-2004
DEFINITION RPAMMSEQ0014829 Roche Palo Alto Mus musculus STS genomic, sequence
tagged site.
ACCESSION BV099496

```
VERSION BV099496.1 GI:40812576
KEYWORDS STS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.
1 (bases 1 to 398)
USUKA,J., LIAO,G., Cheng,J., Nguyen,A., Bach,C., Puech,A.,
McPherson,J.D., Foernzler,D. and Peltz,G.
TITLE Mus musculus SNPs
JOURNAL Unpublished (2003)
COMMENT
Contact: Jonathan Usuka
Roche Palo Alto Genetics and Genomics Department
3431 Hillview Ave, Mailstop S3-1, Palo Alto, CA 94024, USA
Tel: 6508555807
Email: Jonathan.Usuka@roche.com
Primer A: No primer submitted with this STS
Primer B: No primer submitted with this STS.
Location/Qualifiers
FEATURES
source 1..398
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/map="1-13610-13406-CAAA01163581.1.1.19206"
/notes="SNPs developed from assay sequences derived from 15
different strains of mice (as of October 1, 2003). Those
strains include A/J, A/HeJ, 129/Sv, AKR/J, B10.D2-H2/cSnJ,
BALB/cByJ, BALB/cJ, C3H/HeJ, C57BL/6J, CAST/Ei, DBA/2J,
MRL/MpJ, NZB/BinJ, NZW/Lac, SPRET/Ei."
<1..398
STS
ORIGIN
Query Match 87.2%; Score 21.8; DB 10; Length 398;
Best Local Similarity 92.0%; Pred. No. 15;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 AAAAAAAAAATCAGCGCCAGGCATGG 25
|||||
Db 168 AAAAAAAAAATCTTGGCCAGGCATGG 144
|||||
RESULT 23
AF270524/c 701 bp DNA linear PRI 15-DEC-2000
LOCUS Homo sapiens clone 2002c9t3_SMARCA5/SHGC-33076 map 13qtel sequence.
DEFINITION AF270524
ACCESSION AF270524
VERSION AF270524.2 GI:11863208
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 701)
REFERENCE Riethman,H.C. and Moyzis,R.K.
AUTHORS Integration of telomeric DNA sequences with the human reference
TITLE sequence
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 701)
AUTHORS Riethman,H.C. and Moyzis,R.K.
TITLE Direct Submission
JOURNAL Submitted (19-MAY-2000) Molecular Genetics, The Wistar Institute,
3601 Spruce St., Philadelphia, PA 19104, USA
REFERENCE 3 (bases 1 to 701)
AUTHORS Riethman,H.C. and Moyzis,R.K.
TITLE Direct Submission
JOURNAL Submitted (15-DEC-2000) Molecular Genetics, The Wistar Institute,
3601 Spruce St., Philadelphia, PA 19104, USA
REMARK Sequence update by submitter
COMMENT On Dec 15, 2000 this sequence version replaced gi:9755066.
FEATURES
source 1..701
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="13"
/map="13qtel"
/clone="2002c9t3_SMARCA5/SHGC-33076"
ORIGIN
Query Match 87.2%; Score 21.8; DB 8; Length 701;
Best Local Similarity 92.0%; Pred. No. 15;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 AAAAAAAAAATCAGCGCCAGGCATGG 25
|||||
Db 60 AAAAAAAAAATCTTGGCCAGGCATGG 36
|||||
RESULT 24
BV69872/c
```

```
VERSION BV159481/c
KEYWORDS STS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.
1 (bases 1 to 398)
USUKA,J., LIAO,G., Cheng,J., Nguyen,A., Bach,C., Puech,A.,
McPherson,J.D., Foernzler,D. and Peltz,G.
TITLE Mus musculus SNPs
JOURNAL Unpublished (2003)
COMMENT
Contact: Jonathan Usuka
Roche Palo Alto Genetics and Genomics Department
3431 Hillview Ave, Mailstop S3-1, Palo Alto, CA 94024, USA
Tel: 6508555807
Email: Jonathan.Usuka@roche.com
Primer A: No primer submitted with this STS
Primer B: No primer submitted with this STS.
Location/Qualifiers
FEATURES
source 1..398
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/map="1-13610-13406-CAAA01163581.1.1.19206"
/notes="SNPs developed from assay sequences derived from 15
different strains of mice (as of October 1, 2003). Those
strains include A/J, A/HeJ, 129/Sv, AKR/J, B10.D2-H2/cSnJ,
BALB/cByJ, BALB/cJ, C3H/HeJ, C57BL/6J, CAST/Ei, DBA/2J,
MRL/MpJ, NZB/BinJ, NZW/Lac, SPRET/Ei."
<1..398
STS
ORIGIN
Query Match 87.2%; Score 21.8; DB 10; Length 398;
Best Local Similarity 92.0%; Pred. No. 15;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 AAAAAAAAAATCAGCGCCAGGCATGG 25
|||||
Db 168 AAAAAAAAAATCTTGGCCAGGCATGG 144
|||||
RESULT 22
BV159481/c 398 bp DNA linear STS 14-MAY-2004
LOCUS R6AMSE0035768 Roche Palo Alto Mus musculus STS genomic, sequence
DEFINITION tagged site.
ACCESSION BV159481
VERSION BV159481.1 GI:47174716
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.
1 (bases 1 to 398)
USUKA,J., LIAO,G., Cheng,J., Nguyen,A., Bach,C., Puech,A.,
McPherson,J.D., Foernzler,D. and Peltz,G.
TITLE Mus musculus SNPs
JOURNAL Unpublished (2003)
COMMENT
Contact: Jonathan Usuka
Roche Palo Alto Genetics and Genomics Department
3431 Hillview Ave, Mailstop S3-1, Palo Alto, CA 94024, USA
Tel: 6508555807
Email: Jonathan.Usuka@roche.com
Primer A: No primer submitted with this STS
Primer B: No primer submitted with this STS.
Location/Qualifiers
FEATURES
```

Matches	23;	Conservative	0;	Mismatches	2;	Indels	0;	Gaps	0
QY	1	AAAAAAAAATCACGCCAGGCATGG 25							
Db	551	AAAAAAAAAAGCAAGCCAGGCATGG 527							
RESULT 25									
BV590677/c									
LOCUS	BV590677	861 bp	DNA	linear	STS 12-APR-2005				
DEFINITION	G591P643551RG4.T0 Clint Pan troglodytes verus STS genomic, sequence tagged site.								
ACCESSION	BV590677								
VERSION	BV590677.1	GI:62506406							
KEYWORDS	STS.								
SOURCE	Pan troglodytes verus								
ORGANISM	Pan troglodytes verus								
REFERENCE	Pan troglodytes verus								
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Pan.								
TITLE	1 (bases 1 to 861)								
JOURNAL	Mikkelsen,T.S., Hillier,W.L., Eichler,E.E., Zody,M.C. and Jaffe,D.B.								
COMMENT	Initial Sequence of the Chimpanzee Genome and Comparison with the Human Genome Unpublished (2005) Contact: Michael C. Zody Broad Institute of MIT and Harvard 320 Charles Street, Cambridge, MA 02141, USA Tel: 6172580933 Fax: 6172580903 Email: mczody@broad.mit.edu Primer A: No sequence submitted Primer B: No sequence submitted STS size: 861 Protocol: 23,021,928 chimpanzee whole genome shotgun reads were aligned to the Human genome NCBI Build 34 (hg16,July 2003) . Chimp WGS reads were from 9 donors, including Clint (Pan troglodytes verus), 3 other Pan troglodytes verus chimps (Donald,Karlien,Yvonne), 3 Pan troglodytes troglodytes chimps (Noemie,Masuku,Clara) and 2 chimps of unknown origin (Gon,Unknown Chimp) . Common names: Pan troglodytes verus is the western chimp and Pan troglodytes troglodytes is the central chimp. To be included in chimpanzee SNP discovery, a read must be at least 500bp in length, at least 50% of its base calls must have Phred score >= 20, at least 30% of its base calls must satisfy SNQS(30,25)(single strand NQS, the base in question has Phred score >= 30, the surrounding 10 bases in the read have Phred score >= 25), and the read must have at least 200 bp SNQS(30,25) bases. Reads not uniquely placed in the genome and read pairs whose two ends were not consistently placed were discarded. After above filtering, NQS(30,25) standard was applied to all pairs of overlapping reads to call NQS bases and SNPs. Alignments (between two reads) with less than 100 NQS bases or with SNP rate > 0.01 were discarded. To exclude alignment between two copies of a single read, comparisons between two reads that share 95% of their genome alignments (>=95% bases of read A and >=95% bases of read B were placed at the same locus of human genome) were discarded.								
FEATURES	Location/Qualifiers								
source	1..861 /organism="Pan troglodytes verus"								

LOCUS	BV669872	749 bp	DNA	linear	STS 16-APR-2005
DEFINITION	S216P62163RD9.T0 Masuku Pan troglodytes troglodytes STS genomic, sequence tagged site.				
ACCESSION	BV669872				
VERSION	BV669872.1	GI:62697869			
KEYWORDS	STS.				
SOURCE	Pan troglodytes troglodytes				
ORGANISM	Pan troglodytes troglodytes				
REFERENCE	Pan troglodytes troglodytes				
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Pan.				
TITLE	1 (bases 1 to 749)				
JOURNAL	Mikkelsen,T.S., Hillier,W.L., Eichler,E.E., Zody,M.C. and Jaffe,D.B.				
COMMENT	Initial Sequence of the Chimpanzee Genome and Comparison with the Human Genome Unpublished (2005) Contact: Michael C. Zody Broad Institute of MIT and Harvard 320 Charles Street, Cambridge, MA 02141, USA Tel: 6172580933 Fax: 6172580903 Email: mczody@broad.mit.edu Primer A: No sequence submitted Primer B: No sequence submitted STS size: 749 Protocol: 23,021,928 chimpanzee whole genome shotgun reads were aligned to the Human genome NCBI Build 34 (hg16,July 2003) . Chimp WGS reads were from 9 donors, including Clint (Pan troglodytes verus), 3 other Pan troglodytes verus chimps (Donald,Karlien,Yvonne), 3 Pan troglodytes troglodytes chimps (Noemie,Masuku,Clara) and 2 chimps of unknown origin (Gon,Unknown Chimp) . Common names: Pan troglodytes verus is the western chimp and Pan troglodytes troglodytes is the central chimp. To be included in chimpanzee SNP discovery, a read must be at least 500bp in length, at least 50% of its base calls must have Phred score >= 20, at least 30% of its base calls must satisfy SNQS(30,25)(single strand NQS, the base in question has Phred score >= 30, the surrounding 10 bases in the read have Phred score >= 25), and the read must have at least 200 bp SNQS(30,25) bases. Reads not uniquely placed in the genome and read pairs whose two ends were not consistently placed were discarded. After above filtering, NQS(30,25) standard was applied to all pairs of overlapping reads to call NQS bases and SNPs. Alignments (between two reads) with less than 100 NQS bases or with SNP rate > 0.01 were discarded. To exclude alignment between two copies of a single read, comparisons between two reads that share 95% of their genome alignments (>=95% bases of read A and >=95% bases of read B were placed at the same locus of human genome) were discarded.				
FEATURES	Location/Qualifiers				
source	1..749 /organism="Pan troglodytes troglodytes" /mol_type="genomic DNA" /sub_species="troglodytes" /db_xref="taxon:37011" /clone_lib="Masuku" <1..>749				
STS					
ORIGIN					
Query Match	87.2%;	Score 21.8;	DB 10;	Length 749;	
Best Local Similarity	92.0%;	Pred. No. 15;			

/mol_type="genomic DNA"
/sub_species="verus"
/db_xref="taxon:37012"
/clone_lib="Clint"
<1..5661

STS ORIGIN

Query Match 87.2%; Score 21.8; DB 10; Length 861;
Best Local Similarity 92.0%; Pred. No. 15;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AAAAAAAAAATCAGCGCCAGCATGG 25
|||||
788 AAAAAAAAAATCAGCGCCAGCATGG 764

RESULT 26
BV594288 942 bp DNA linear STS 12-APR-2005
LOCUS sxj38a02.b1 Clint Pan troglodytes verus STS genomic, sequence
DEFINITION tagged site.

ACCESSION BV594288
VERSION BV594288.1 GI:62510017
KEYWORDS STS.
SOURCE Pan troglodytes verus
ORGANISM Pan troglodytes verus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Pan.

REFERENCE 1 (bases 1 to 942)
Mikkelsen,T.S., Hillier,W.L., Eichler,E.E., Zody,M.C. and
Jaffe,D.B.

TITLE Initial Sequence of the Chimpanzee Genome and Comparison with the
Human Genome
JOURNAL Unpublished (2005)
COMMENT

Contact: Michael C. Zody
Broad Institute of MIT and Harvard
320 Charles Street, Cambridge, MA 02141, USA
Tel: 6172580933

Fax: 6172580903

Email: mczody@broad.mit.edu

Primer A: No sequence submitted

Primer B: No sequence submitted

STS size: 942

Protocol:

23,021,928 chimpanzee whole genome shotgun reads were aligned to
the Human genome NCBI

Build 34 (hg16,July 2003). Chimp WGS reads were from 9 donors,

including Clint (Pan

troglodytes verus), 3 other Pan troglodytes verus chimps

(Donald,Karlien,Yvonne), 3 Pan

troglodytes troglodytes chimps (Noemie,Masuku,Clara) and 2 chimps

of unknown origin

(Gon Unknown Chimp). Common names: Pan troglodytes verus is the

western chimp and Pan

troglodytes troglodytes is the central chimp. To be included in

chimpanzee SNP discovery, a

read must be at least 500bp in length, at least 50% of its base

calls must have phred

score >= 20, at least 30% of its base calls must satisfy

SNQS(30,25)(single strand NQS, the

base in question has phred score >= 30, the surrounding 10 bases in

the read have phred

score >= 25), and the read must have at least 200 bp SNQS(30,25)

placed in the genome and read pairs whose two ends were not

consistently placed were

discarded. After above filtering, NQS(30,25) standard was applied

to all pairs of

overlapping reads to call NQS bases and SNPs. Alignments (between

two reads) with less

than 100 NQS bases or with SNP rate > 0.01 were discarded. To

exclude alignment between two
copies of a single read, comparisons between two reads that share
95% of their genome
alignments (>=95% bases of read A and >=95% bases of read B were
placed at the same locus
of human genome) were discarded.
Location/Qualifiers

FEATURES source

1..942
/organism="Pan troglodytes verus"
/mol_type="genomic DNA"
/sub_species="verus"
/db_xref="taxon:37012"
/clone_lib="Clint"
<1..942

STS ORIGIN

Query Match 87.2%; Score 21.8; DB 10; Length 942;
Best Local Similarity 92.0%; Pred. No. 15;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AAAAAAAAAATCAGCGCCAGCATGG 25

|||||
Db 106 AAAAAAAAAATCAGCGCGGCATGG 130

RESULT 27

AK090515/c

LOCUS Homo sapiens cDNA FLJ33196 fis, clone ADRGL2006034.
DEFINITION AK090515
ACCESSION AK090515
VERSION AK090515.1 GI:21748691

KEYWORDS oligo capping; fis (full insert sequence).

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

REFERENCE

1

AUTHORS

Ota,T., Suzuki,Y., Nishikawa,T., Otsuki,T., Sugiyama,T., Irie,R.,
Wakanatsu,A., Hayashi,K., Sato,H., Nagai,K., Kimura,K., Makita,H.,
Sekine,M., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T.,
Ishii,S., Yamamoto,J., Saito,K., Kawai,Y., Isono,Y., Nakamura,Y.,
Nagahari,K., Murakami,K., Yasuda,T., Iwayanagi,T., Wagatsuma,M.,
Shiratori,A., Sudo,H., Hosoiri,T., Kaku,Y., Kodaira,H., Kondo,H.,
Sugawara,M., Takahashi,M., Kanda,K., Yokoi,T., Furuya,T.,
Kikkawa,S., Omura,Y., Abe,K., Kamiyama,K., Katsuta,N., Sato,K.,
Tanikawa,M., Yamazaki,M., Ninomiya,K., Ishibashi,T., Yamashita,H.,
Murakawa,K., Fujimori,K., Tanai,H., Kimata,M., Watanabe,M.,
Hiraoka,S., Chiba,Y., Ishida,S., Ono,Y., Takiguchi,S., Watanabe,S.,
Yosida,M., Hotuta,T., Kusano,J., Kanehori,K., Takahashi-Fujii,A.,
Hara,H., Tanase,T., Nomura,Y., Togiya,S., Komai,F., Hara,R.,
Takeuchi,K., Arita,M., Imose,N., Musashino,K., Yuuki,H., Oshima,A.,
Sasaki,N., Aotsuka,S., Yoshikawa,Y., Matsunawa,H., Ichihara,T.,
Shiohata,N., Sano,S., Moriya,S., Momiyama,H., Satoh,N., Takami,S.,
Terashima,Y., Suzuki,O., Nakagawa,S., Senoh,A., Mizoguchi,H.,
Goto,Y., Shimizu,F., Wakebe,H., Hishigaki,H., Watanabe,T.,
Sugiyama,A., Takemoto,M., Kawakami,B., Yamazaki,M., Watanabe,K.,
Kumegaki,A., Itakura,S., Fukuzumi,Y., Fujimori,Y., Komiyama,M.,
Tashiro,H., Tanigami,A., Fujiwara,T., Ono,T., Yamada,K., Fujii,Y.,
Ozaki,K., Hirao,M., Ohmori,Y., Kawabata,A., Hikiji,T., Kobatake,N.,
Inagaki,H., Ikema,Y., Okamoto,S., Okitani,R., Kawakami,T.,
Noguchi,S., Itoh,T., Shigetani,K., Senba,T., Matsumura,K.,
Nakajima,Y., Mizuno,T., Morinaga,M., Sasaki,M., Togashi,T.,
Oyama,M., Hata,H., Watanabe,M., Komatsu,T., Mizushima-Sugano,J.,
Satoh,T., Shirai,Y., Takahashi,Y., Nakagawa,K., Okumura,K.,
Nagase,T., Nomura,N., Kikuchi,H., Masuho,Y., Yamashita,K.,
Nakai,K., Yada,T., Nakamura,Y., Ohara,O., Isogai,T. and Sugano,S.

Complete sequencing and characterization of 21,243 full-length

human cDNAs

Nat. Genet. 36 (1), 40-45 (2004)

JOURNAL

PUBLISHED

REFERENCE

2

AUTHORS

Nishi,T., Ota,T., Nakagawa,S., Senoh,A., Mizuguchi,H., Inagaki,H.,

Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K., Wagatsuma, M., Murakawa, K., Kanehori, K., Takahashi-Fujii, A., Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S., Nagahari, K., Masuho, Y., Nagai, K. and Isogai, T.
NEDO human cDNA sequencing project

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

Unpublished
3 (bases 1 to 2141)

Isogai, T. and Yamamoto, J.

Direct Submission

Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team) ; 2-6-7
Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan

NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction: Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: HRI and
RAB; annotation: HRI and RAB.

FEATURES
Source

Location/Qualifiers
1..2141
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="ADRG1206034"
/tissue_type="adrenal gland"
/clone_lib="ADRL2"
/note="cloning vector: pME18SFL3"

ORIGIN

Query Match 87.2%; Score 21.8; DB 8; Length 2141;
Best Local Similarity 92.0%; Pred. No. 15;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAAAAAATCAGCGCCAGGCATGG 25

Db 1169 AAAAAATAATCAGCGCCAGGCACGG 1145

RESULT 28
AK097083/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS

AK097083
Homo sapiens cDNA FLJ39764 fis, clone SPLEN2000143.
PRI 30-JAN-2004

AK097083
Homo sapiens

AK097083.1 GI:21756734

oligo capping; fis (full insert sequence).

Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.

Ota, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R.,
Wakamatsu, A., Hayashi, K., Sato, H., Nagai, K., Kimura, K., Makita, H.,
Sekine, M., Obayashi, M., Nishi, T., Shibahara, T., Tanaka, T.,
Ishii, S., Yamamoto, J., Saito, K., Kawai, Y., Isono, Y., Nakamura, Y.,
Nagahari, K., Murakami, K., Yasuda, T., Iwayanagi, T., Wagatsuma, M.,
Shiratori, A., Sudo, H., Hosoiri, T., Kaku, Y., Kodaira, H., Kondo, H.,
Sugawara, M., Takahashi, M., Kanda, K., Yokoi, T., Furuya, T.,
Kikawa, S., Omura, Y., Abe, K., Kamihara, K., Katsuta, N., Sato, K.,
Tanikawa, M., Yamazaki, M., Ninomiya, K., Ishibashi, T., Yamashita, H.,
Murakawa, K., Fujimori, K., Tanai, H., Kimata, M., Watanabe, M.,
Hiraoka, S., Chiba, Y., Ishida, S., Ono, Y., Takiguchi, S., Watanabe, S.,
Yosida, M., Hotuta, T., Kusano, J., Kanehori, K., Takahashi-Fujii, A.,
Hara, H., Tanase, T., Nomura, Y., Togiya, S., Komai, F., Hara, R.,
Takeuchi, K., Arita, M., Inose, N., Musashino, K., Yuuki, H., Oshima, A.,
Sasaki, N., Aotsuka, S., Yoshikawa, Y., Matsunawa, H., Ichihara, T.,
Shiohata, N., Sano, S., Moriya, S., Momiyama, H., Satoh, N., Takami, S.,
Teraashina, Y., Suzuki, O., Nakagawa, S., Senoh, A., Mizoguchi, H.,
Goto, Y., Shimizu, F., Wakebe, H., Hishigaki, H., Watanabe, T.,

Sugiyama, A., Takemoto, M., Kawakami, B., Yamazaki, M., Watanabe, K.,
Kunagai, A., Itakura, S., Fukuzumi, Y., Fujimori, Y., Komiyama, M.,
Tashiro, H., Tanigami, A., Fujiwara, T., Ono, T., Yamada, K., Fujii, Y.,
Ozaki, K., Hirao, M., Ohmori, Y., Kawabata, A., Hikiiji, T., Kobatake, N.,
Inagaki, H., Ikema, Y., Okamoto, S., Okitani, R., Kawakami, T.,
Noguchi, S., Itoh, T., Shigetani, K., Senba, T., Matsumura, K.,
Nakajima, Y., Mizuno, T., Morinaga, M., Sasaki, M., Togaashi, T.,
Oyama, M., Hata, H., Watanabe, M., Komatsu, T., Mizushima-Sugano, J.,
Satoh, T., Shirai, Y., Takahashi, Y., Nakagawa, K., Okumura, K.,
Nagase, T., Nomura, N., Kikuchi, H., Masuho, Y., Yamashita, R.,
Nakai, K., Yada, T., Nakamura, Y., Ohara, O., Isogai, T. and Sugano, S.
Complete sequencing and characterization of 21,243 full-length
human cDNAs

Nat. Genet. 36 (1), 40-45 (2004)

14702039

JOURNAL
PUBMED
REFERENCE
AUTHORS

Nishi, T., Ota, T., Nakagawa, S., Senoh, A., Mizuguchi, H., Inagaki, H.,
Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S.,
Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T.,
Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M.,
Kikuchi, H., Kanda, K., Wagatsuma, M., Murakawa, K., Kanehori, K.,
Takahashi-Fujii, A., Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y.,
Sugano, S., Nagahari, K., Masuho, Y., Nagai, K. and Isogai, T.
NEDO human cDNA sequencing project

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

Unpublished
3 (bases 1 to 2530)

Isogai, T. and Yamamoto, J.

Direct Submission

Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team) ; 2-6-7
Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction: Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: HRI and
RAB; annotation: HRI and RAB.

FEATURES
source

Location/Qualifiers
1..2530
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="SPLEN2000143"
/tissue_type="spleen"
/clone_lib="SPLEN2"
/note="cloning vector: pME18SFL3"

ORIGIN

Query Match 87.2%; Score 21.8; DB 8; Length 2530;
Best Local Similarity 92.0%; Pred. No. 15;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAAAAAATCAGCGCCAGGCATGG 25

Db 1135 AAAAAATAATCAGCGCCAGGCACGG 1111

RESULT 29

S81092/c

LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE

S81092
acyl-coenzyme A:cholesterol acyltransferase [mice, peritoneal
macrophages, mRNA, 3041 nt].
S81092
S81092.1 GI:1478335
Mus sp.
Mus sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 3041)

3041 bp mRNA linear ROD 03-AUG-1996

----- Genome Center
 Center: Wellcome Trust Sanger Institute
 Center code: SC
 Web site: <http://www.sanger.ac.uk>
 Contact: vega@sanger.ac.uk

This sequence was finished as follows unless otherwise noted: all chemistry were either double-stranded or sequenced with an alternate 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.
 WI2-801302G5 is from a Whitehead human fosmid library VECTOR: pEpifos-5 Eco721.

FEATURES

source

Location/Qualifiers

1..35616
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="1"
 /clone="WI2-801302G5"
 /clone_lib="Whitehead_fosmid"
 2000
 /note="Clone_right_end: RP11-5P18"

misc_feature

ORIGIN

Query Match 87.2%; Score 21.8; DB 8; Length 35616;
 Best Local Similarity 92.0%; Pred. No. 15;
 Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 AAAAAAAAAATCAGCGCCAGCGCATGG 25

Db 10247 AAAAAAAAAAGCAAGCGCCAGCGCATGG 10223

RESULT 32

AC104230

LOCUS Homo sapiens chromosome 8 clone RP11-351N23 map 8, LOW-PASS
 DEFINITION SEQUENCE SAMPLING.

ACCESSION AC104230

VERSION AC104230.1 GI:17386398

KEYWORDS HTG; HTGS_PHASE0.

SOURCE Homo sapiens

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Homo.

REFERENCE 1 (bases 1 to 46784)

AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,B.

TITLE Homo sapiens chromosome 8, clone RP11-351N23

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 46784)

AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,B., Ali,A., Allen,N.,
 Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhgalter,B.,
 Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
 Choepel,Y., Collangelo,M., Collins,S., Collamore,A., Cook,A.,
 Cooke,P., DeAtellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
 Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S.,
 Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
 Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliiev,I., Johnson,R.,
 Jones,C., Kamat,A., Karatas,A., Kells,C., Lakocque,K.,
 Lamazares,R., Landers,T., Lehotzky,J., Levine,R., Liu,G.,
 Maclean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
 McCarty,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J.,
 Meneus,L., Mihova,T., Mianga,V., Murphy,T., Naylor,J., Nguyen,C.,
 Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,B.,
 Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
 Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
 Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupbach,R.,
 Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
 Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,

Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
 Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
 Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
 Direct Submission
 Submitted (06-DEC-2001) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>
 ----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L21645

Center clone name: 351_N_23

* NOTE: This record contains 57 individual
 * sequencing reads that have not been assembled into
 * contigs. Runs of N are used to separate the reads
 * and the order in which they appear is completely
 * arbitrary. Low-pass sequence sampling is useful for
 * identifying clones that may be gene-rich and allows
 * overlap relationships among clones to be deduced.
 * However, it should not be assumed that this clone
 * will be sequenced to completion. In the event that
 * the record is updated, the accession number will
 * be preserved.

* 1 699: contig of 699 bp in length

* 799: gap of 100 bp

* 800 1509: contig of 710 bp in length

* 1510 1509: gap of 100 bp

* 1610 2310: contig of 701 bp in length

* 2411 3155: contig of 745 bp in length

* 3156 3255: gap of 100 bp

* 3256 3986: contig of 731 bp in length

* 3987 4086: gap of 100 bp

* 4087 4806: contig of 720 bp in length

* 4807 4906: gap of 100 bp

* 4907 5644: contig of 738 bp in length

* 5645 5744: gap of 100 bp

* 5745 6461: contig of 717 bp in length

* 6462 6561: gap of 100 bp

* 6562 7280: contig of 719 bp in length

* 7281 7380: gap of 100 bp

* 7381 8096: contig of 716 bp in length

* 8097 8196: gap of 100 bp

* 8197 8921: contig of 725 bp in length

* 8922 9021: gap of 100 bp

* 9022 9736: contig of 715 bp in length

* 9737 9836: gap of 100 bp

* 9837 10571: contig of 735 bp in length

* 10572 10671: gap of 100 bp

* 10672 11412: contig of 741 bp in length

* 11413 11512: gap of 100 bp

* 11513 12355: contig of 743 bp in length

* 12356 13067: contig of 100 bp

* 13068 13167: contig of 712 bp in length

* 13168 13913: contig of 100 bp

* 13914 14013: gap of 100 bp

* 14014 14755: contig of 741 bp in length

* 14755 14854: gap of 100 bp

* 14855 15554: contig of 700 bp in length

* 15555 15654: gap of 100 bp

* 15655 16368: contig of 714 bp in length

* 16369 16468: gap of 100 bp

* 16469 17309: contig of 741 bp in length

* 17310 17309: gap of 100 bp

* 17310 18055: contig of 746 bp in length

* 18056 18155: gap of 100 bp

TITLE
 JOURNAL

COMMENT

/db_xref="taxon:9606"
/chromosome="8"
/map="8"
/clone="RP11-351N23"
/clone_lib="RPCI-11 Human Male BAC"
700..799
/estimated_length=100
1510..1609
/estimated_length=100
2311..2410
/estimated_length=100
3156..3255
/estimated_length=100
3987..4086
/estimated_length=100
4807..4906
/estimated_length=100
5645..5744
/estimated_length=100

Query Match 87.2%; Score 21.8; DB 14; Length 46784;
Best Local Similarity 92.0%; Pred. No. 15;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAAAAAATCAGCGCCAGGCATGG 25
|||||||
DB 33028 AAAAAAAAAATCAGCGCCAGGCATGG 33052

RESULT 33
AL358393
LOCUS
DEFINITION
AL358393 50175 bp DNA linear PRI 18-MAY-2005
Human DNA sequence from clone Rp11-540K16 on chromosome 1 Contains
the 3' end of a novel gene (LOC284646) and a novel gene, complete
sequence.
AL358393
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
1 (bases 1 to 50175)
Almeida, J.
Direct Submission
Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vegas@sanger.ac.uk
Clone requests: clonerequest@sanger.ac.uk
On Nov 16, 2001 this sequence version replaced gi:15722103.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em., EMBL; Sw., SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr1
Rp11-540K16 is from the library RPCI-11.2 constructed by the group
of Pieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VBCIOR: pBACe3.6
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: vegas@sanger.ac.uk

This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such

18156 18877: contig of 722 bp in length
* 18878 18977: gap of 100 bp in length
* 18978 19707: contig of 730 bp in length
* 19708 19807: gap of 100 bp in length
* 19808 20545: contig of 738 bp in length
* 20546 21351: contig of 100 bp in length
* 21352 21451: gap of 100 bp in length
* 21452 22171: contig of 720 bp in length
* 22172 22271: gap of 100 bp in length
* 22272 22989: contig of 718 bp in length
* 22990 23089: gap of 100 bp in length
* 23090 23801: contig of 712 bp in length
* 23802 23901: gap of 100 bp in length
* 23902 24637: contig of 736 bp in length
* 24638 24737: gap of 100 bp in length
* 24738 25495: contig of 758 bp in length
* 25496 25595: gap of 100 bp in length
* 25596 26286: contig of 691 bp in length
* 26287 26386: gap of 100 bp in length
* 26387 27116: contig of 730 bp in length
* 27117 27216: gap of 100 bp in length
* 27217 27919: contig of 703 bp in length
* 27920 28019: gap of 100 bp in length
* 28020 28767: contig of 748 bp in length
* 28768 28867: gap of 100 bp in length
* 28868 29589: contig of 722 bp in length
* 29590 30402: contig of 713 bp in length
* 30403 30502: gap of 100 bp in length
* 30503 31182: contig of 680 bp in length
* 31183 31282: gap of 100 bp in length
* 31283 31977: contig of 695 bp in length
* 31978 32077: gap of 100 bp in length
* 32078 32772: contig of 695 bp in length
* 32773 32872: gap of 100 bp in length
* 32873 33586: contig of 714 bp in length
* 33587 33686: gap of 100 bp in length
* 33687 34430: contig of 744 bp in length
* 34431 34530: gap of 100 bp in length
* 34531 35269: contig of 739 bp in length
* 35270 35369: gap of 100 bp in length
* 35370 36086: contig of 717 bp in length
* 36087 36186: gap of 100 bp in length
* 36187 36896: contig of 710 bp in length
* 36897 36996: gap of 100 bp in length
* 36997 37707: contig of 711 bp in length
* 37708 37807: gap of 100 bp in length
* 37808 38559: contig of 752 bp in length
* 38560 38659: gap of 100 bp in length
* 38660 39352: contig of 693 bp in length
* 39353 39452: gap of 100 bp in length
* 39453 40153: contig of 701 bp in length
* 40154 40253: gap of 100 bp in length
* 40254 40970: contig of 717 bp in length
* 40971 41070: gap of 100 bp in length
* 41071 41819: contig of 749 bp in length
* 41820 41919: gap of 100 bp in length
* 41920 42645: contig of 726 bp in length
* 42646 42745: gap of 100 bp in length
* 42746 43443: contig of 698 bp in length
* 43444 43544: gap of 100 bp in length
* 43544 44272: contig of 729 bp in length
* 44273 44372: gap of 100 bp in length
* 44373 45111: contig of 739 bp in length
* 45112 45211: gap of 100 bp in length
* 45212 45962: contig of 751 bp in length
* 45963 46062: gap of 100 bp in length
* 46063 46784: contig of 722 bp in length.

Location/Qualifiers
1. 46784
/organism="Homo sapiens"
/mol_type="genomic DNA"

FEATURES
source

as compressions and repeats; all regions were covered by at least one subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

```

FEATURES
    source
        Location/Qualifiers
            1..50175
                /organism="Homo sapiens"
                /mol_type="genomic DNA"
                /db_xref="taxon:9606"
                /chromosome="1"
                /clone="RP11-540K16"
                /clone_lib="RPC1-11.2"
                2000
                /note="Clone right end: RP11-309G3"
                join(AL356267.27:177684..177775,
                AL356267.27:178537..178666,5281..5600)
                /gene="RP11-540K16.1"
                /locus_tag="RP11-540K16.1-001"
                join(AL356267.27:177684..177775,
                AL356267.27:178537..178666,5281..5600)
                /gene="RP11-540K16.1"
                /locus_tag="RP11-540K16.1-001"
                /product="novel transcript"
                /note="match=ESTs: Em:BM551928.1
                match: CDNAS: Em:BC039493.1"
                join(AL356267.27:178537..178633,
                AL356267.27:178910..179034,5281..5599)
                /gene="RP11-540K16.1"
                /locus_tag="RP11-540K16.1-002"
                join(AL356267.27:178537..178633,
                AL356267.27:178910..179034,5281..5599)
                /gene="RP11-540K16.1"
                /locus_tag="RP11-540K16.1-002"
                /product="novel transcript"
                /note="match=ESTs: Em:AL680706.1"
                5575..5580
                /gene="RP11-540K16.1"
                /locus_tag="RP11-540K16.1-002"
                5600
                /gene="RP11-540K16.1"
                /locus_tag="RP11-540K16.1-001"
                complement(join(13863..14200,14348..14541))
                /locus_tag="RP11-540K16.2-001"
                complement(join(13863..14200,14348..14541))
                /locus_tag="RP11-540K16.2-001"
                /product="putative novel transcript"
                /note="match=ESTs: Em:AW204111.1 Em:BB550132.1"
                complement(13863)
                /locus_tag="RP11-540K16.2-001"
                complement(13884..13889)
                /locus_tag="RP11-540K16.2-001"
                48176
                /note="Clone_left_end: RP11-380B22"

ORIGIN
Query Match      87.2%; Score 21.8; DB 8; Length 50175;
Best Local Similarity 92.0%; Pred. No. 15;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAAAAAACAGCGCAGGCATGG 25
    ||||| ||| ||||| ||||| |||||
Db 37635 AAAAAAAAAACAGCGCAGGCATGG 37659

RESULT 34
AC104800/c
LOCUS AC104800 Homo sapiens BAC clone RP11-350L15 from 4, complete sequence.
DEFINITION AC104800
ACCESSION AC104800.5 GI:21263351
VERSION AC104800.5
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 54618)
Sulston,J.B. and Waterston,R.
Toward a complete human genome sequence
Genome Res. 8 (11), 1097-1108 (1998)
9847074

2 (bases 1 to 54618)
Tomlinson,C., Meyer,R. and Creason,K.
The sequence of Homo sapiens BAC clone RP11-350L15
Unpublished (2001)

3 (bases 1 to 54618)
Waterston,R.H.
Direct Submission

Submitted (21-DEC-2001) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA

4 (bases 1 to 54618)
Waterston,R.H.
Direct Submission

Submitted (01-MAY-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA

5 (bases 1 to 54618)
Waterston,R.H.
Direct Submission

Submitted (30-MAY-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA

6 (bases 1 to 54618)
Waterston,R.
Direct Submission

Submitted (12-JUN-2002) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On May 30, 2002 this sequence version replaced gi:20377064.

----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu/gsc>
Contact: sapiens@watson.wustl.edu
----- Summary Statistics

Center project name: H_NH0350L15

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:
The RPC1-11 human BAC library was made from the blood of one male donor, as described by Osoegawa,K., Woon,P.Y., Zhao,B., Frengen,B., Tateno,M., Cataneese,J.J. and de Jong,P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>
VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-357L18, 2000 bp overlap;
 the clone sequenced to the right is RP11-568024, 2000 bp overlap.
 Actual start of this clone is at base position 9509 of RP11-357L18;
 actual end is at base position 77842 of RP11-568024.

FEATURES

```

source
  1. .54618
    /organism="Homo sapiens"
    /mol_type="genomic DNA"
    /db_xref="taxon:9606"
    /chromosome="4"
    /map="4"
    /clone="RP11-350L15"
    /clone_lib="RPCi-11"
    /rpt_family="L2"
  30. .251
    /rpt_family="L2"
  369. .497
    /rpt_family="MIR"
  812. .1108
    /rpt_family="Alu"
  1332. .1767
    /rpt_family="ERV1"
  2574. .2655
    /rpt_family="L1"
  2656. .2685
    /rpt_family="L1"
  2690. .3097
    /rpt_family="AT_rich"
    /rpt_family="L1"
  3086. .3173
    /rpt_family="L1"
  3141. .3564
    /rpt_family="L1"
  3565. .5771
    /rpt_family="L1"
  5749. .5940
    /rpt_family="L1"
  6342. .6465
    /rpt_family="MaLR"
  7644. .7655
    /rpt_family="L1"
  7656. .8022
    /rpt_family="L1"
  8023. .8286
    /rpt_family="L1"
  8287. .8313
    /rpt_family="(TTG)n"
  8314. .8953
    /rpt_family="L1"
  9743. .9823
    /rpt_family="MaLR"
  10794. .10837
    /rpt_family="AT_rich"
  11545. .11873
    /rpt_family="L1"
  11900. .12152
    /rpt_family="Alu"
  12376. .12423
    /rpt_family="(TATG)n"
  12484. .12550
    /rpt_family="L1"
  12551. .12575
    /rpt_family="(CAAAA)n"
  12576. .13165
    /rpt_family="L1"
  13184. .13224
    /rpt_family="L1"
  13290. .13448
    /rpt_family="Alu"
  13506. .13531
    /rpt_family="AT_rich"
  15831. .16294
    /rpt_family="L1"
  16290. .16489

```

```

repeat_region      /rpt_family="L1"
16568. .16735
repeat_region      /rpt_family="(TA)n"
16731. .16911
repeat_region      /rpt_family="(TTATA)n"
16925. .17098
repeat_region      /rpt_family="(TTATA)n"
17099. .17200
repeat_region      /rpt_family="(TA)n"
17204. .17654
repeat_region      /rpt_family="(TTATA)n"
17655. .18080
repeat_region      /rpt_family="(TATAA)n"
18086. .18245
repeat_region      /rpt_family="(TTATA)n"
18257. .18488
repeat_region      /rpt_family="(TTATA)n"
18493. .18558
repeat_region      /rpt_family="(TA)n"
18603. .18732
repeat_region      /rpt_family="(TA)n"
18765. .19018
repeat_region      /rpt_family="L1"
19548. .20379
repeat_region      /rpt_family="L1"
20463. .20486
repeat_region      /rpt_family="AT_rich"
20502. .20527
repeat_region      /rpt_family="AT_rich"
20642. .20820
repeat_region      /rpt_family="L1"
20950. .21613
repeat_region      /rpt_family="L1"
21614. .21923
repeat_region      /rpt_family="Alu"
21924. .22811
repeat_region      /rpt_family="L1"
22785. .23287
repeat_region      /rpt_family="L1"
23304. .23796
repeat_region      /rpt_family="L1"
23797. .24104
repeat_region      /rpt_family="Alu"
24105. .24151
repeat_region      /rpt_family="L1"
24145. .24914
repeat_region

```

Query Match 87.2%; Score 21.8; DB 8; Length 54618;
 Best Local Similarity 92.0%; Pred. No. 15;
 Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AAAAAAAAAATCAGCGCAGCATGG 25
 |||||
 Db 24117 AAAAAAAAAATCTCAGCGCAGCATGG 24093

RESULT 35

AC105044
 LOCUS AC105044 Homo sapiens chromosome 8 clone CTD-3089F7 map 8, linear HTG 22-DEC-2001
 DEFINITION SAMPLING.
 AC105044
 VERSION AC105044.1 GI:17977423
 KEYWORDS HTG; HTGS_PHASE0.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Hominoidea; Homo.
 REFERENCE 1 (bases 1 to 65882)
 AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
 TITLE Homo sapiens chromosome 8, clone CTD-3089F7
 JOURNAL Unpublished

* 40736 40835: gap of 100 bp
* 40836 contig of 746 bp in length
* 41581: gap of 100 bp
* 41582 gap of 100 bp
* 41582 contig of 721 bp in length
* 42402: gap of 100 bp
* 42502: gap of 100 bp
* 42502: contig of 724 bp in length
* 43226: gap of 100 bp
* 43227 43326: gap of 100 bp
* 43327 44032: contig of 706 bp in length
* 44033 44132: gap of 100 bp
* 44133 44881: contig of 749 bp in length
* 44881: gap of 100 bp
* 44882 45717: contig of 736 bp in length
* 45717: gap of 100 bp
* 45718 46566: contig of 749 bp in length
* 46566: gap of 100 bp
* 46567 47406: contig of 740 bp in length
* 47406: gap of 100 bp
* 47407 47506: gap of 100 bp
* 47507 48241: contig of 735 bp in length
* 48241: gap of 100 bp
* 48342 49068: contig of 727 bp in length
* 49068: gap of 100 bp
* 49069 49168: gap of 100 bp
* 49169 49904: contig of 736 bp in length
* 49904: gap of 100 bp
* 50005 50746: contig of 742 bp in length
* 50746: gap of 100 bp
* 50747 51589: contig of 743 bp in length
* 50847 51589: contig of 743 bp in length
* 51590 51689: gap of 100 bp
* 51690 52423: contig of 734 bp in length
* 52424 52523: gap of 100 bp
* 52524 53277: contig of 754 bp in length
* 53277: gap of 100 bp
* 53278 54105: contig of 728 bp in length
* 53378 54205: gap of 100 bp
* 54106 54957: contig of 752 bp in length
* 54206 55057: gap of 100 bp
* 54958 55058: contig of 703 bp in length
* 55058 55760: gap of 100 bp
* 55761 55860: gap of 100 bp
* 55861 56610: contig of 750 bp in length
* 56610 56710: gap of 100 bp
* 56710: gap of 100 bp

Query Match 87.2%; Score 21.8; DB 14; Length 65882;
Best Local Similarity 92.0%; Pred. No. 15;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AAAAAAAAAATCAGCGCCAGCATGG 25
|||||
Db 9384 AAAAAAAAAATCAGCGCCAGCATGG 9408
|||||

RESULT 36
AL589910/c
LOCUS
DEFINITION Human DNA sequence from clone RP11-57L9 on chromosome 6 Contains part of the gene for a novel protein, complete sequence.
ACCESSION AL589910
VERSION AL589910.9 GI:14041767
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 66591)
Garner, P.
Direct Submission
Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vegas@sanger.ac.uk
On May 15, 2001 this sequence version replaced gi:13897415.
The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:
Em, EMBL; Sw, SWISSPROT; Tr, TrEMBL; Wp, WORMPEP; Information on the WORMPEP database can be found at

http://www.sanger.ac.uk/Projects/C.elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping Group. Further information can be found at http://www.sanger.ac.uk/HGP/Chr6
RP11-57L9 is from the library RPCI-11.1 constructed by the group of Pieter de Jong. For further details see http://www.Chori.org/bacpac/home.htm
VECTOR: pBACE3.6
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: vegas@sanger.ac.uk

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.
FEATURES
Location/Qualifiers
source
1..66591
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="6"
/clone="RP11-57L9"
/clone_lib="RPCI-11.1"
misc_feature
1
/note="Clone left end: RP11-57L9"
gene
join(complement(ALI39098..15:434..575), complement(<43428..54443))
/locus_tag="RP11-57L9.1-001"
mRNA
join(complement(ALI39098..15:434..575), complement(54282..54443), complement(50144..50321), complement(<43428..43496))
/locus_tag="RP11-57L9.1-001"
/note="match: ESTg: AW502864 BF356516 BF356518 BF356520 BF356526 BI49952"
CDS
join(complement(ALI39098..15:434..558), complement(54282..54443), complement(50144..50321), complement(<43428..43496))
/locus_tag="RP11-57L9.1-001"
/standard_name="OTTHUMP0000017110"
/codon_start=1
/protein_id="CA116051.1"
/db_xref="GI:55960366"
/translation="MLQAMLRRLVQSVKEKITGAPSLCAEAILLHLEETDENPHNVEFKYLRQHIGNTLGSIMIEEMKEKTSDRNQEGCYDVTVVQVTKRTOESKEYKEMMHQKNIIMIAVVESMINKFEDETRNQERQKIQEKSHSYTDCSDSDSSLSNQSYKFCQKQLQILDLDPGQPKR"
misc_feature
63055
/note="Clone_right_end: RP11-358H7"
misc_feature
66492
/note="Clone_left_end: RP1-276J11"
ORIGIN
Query Match 87.2%; Score 21.8; DB 8; Length 66591;
Best Local Similarity 92.0%; Pred. No. 15;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 AAAAAAAAAATCAGCGCCAGCATGG 25
|||||
Db 45932 AAAAAAAAAATTAAGCCAGCATGG 45908
|||||

RESULT 37
AC133141
LOCUS
DEFINITION Homo sapiens chromosome 3 clone RP11-421F9, complete sequence.
AC133141 70560 bp DNA linear PRI 31-OCT-2002

ACCESSION AC133141 AC011323
VERSION AC133141.2 GI:24431665
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 70560)
AUTHORS Kaul, R.K., Olson, M.V., Zhou, Y., James, R.A., Rouse, G., Wu, Z.,
Saenphimmachak, C., Buckley, D., Kibukawa, M., Raymond, C. and
Haugen, E.D.
TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 70560)
AUTHORS Kaul, R.K., Olson, M.V., Raymond, C. and Haugen, E.D.
TITLE Direct Submission
JOURNAL Submitted (07-SEP-2002) Genome Center, University of Washington,
Box 352145, Seattle, WA 98195, USA
REFERENCE 3 (bases 1 to 70560)
AUTHORS Kaul, R.K., Olson, M.V., Zhou, Y., James, R.A., Rouse, G., Wu, Z.,
Saenphimmachak, C., Buckley, D., Kibukawa, M., Raymond, C. and
Haugen, E.D.
TITLE Direct Submission
JOURNAL Submitted (31-OCT-2002) Genome Center, University of Washington,
Box 352145, Seattle, WA 98195, USA
COMMENT On Oct 31, 2002 this sequence version replaced gi:22758407.
----- Genome Center
Center: University of Washington Genome Center
Center Code: UWGC
Web site: <http://www.genome.washington.edu>
Contact: uwgchtgs@u.washington.edu
Drafting Center: BCM
----- Project Information
Center project name: chr-3
Center clone name: bc0414
bc0414 was derived from the original BAC (RP11-421P9)
using yeast homologous capture technique.
----- Summary Statistics
Sequencing vector: plasmid; 31% of reads
Sequencing vector: M13; L08821; 69% of reads
Chemistry: Dye-terminator Big Dye; 90% of reads
Chemistry: Dye-primer Bodipy; 10% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 70501 bases at least Q40
Consensus quality: 70557 bases at least Q30
Consensus quality: 70560 bases at least Q20
Insert size: 70560; sum-of-contigs
Quality coverage: 6.7x in Q20 bases; sum-of-contigs

Overlapping Sequences:
5': RP11-733H11 (UMGC:bc0541) AC098477, 4145-bp overlap
3': RP11-659P16 (UMGC:bc0520) AC093416, 1862-bp overlap

Sequence Quality Assessment:
This entry has been annotated with sequence quality
estimates computed by the Phrap assembly program.
All manually edited bases have been reduced to quality zero.
Quality levels above 40 are expected to have less than
1 error in 10,000 bp.
Base-by-base quality values are not generally visible from the
GenBank flat file format but are available as part
of this entry's ASN.1 file.

This sequence was finished as follows unless otherwise noted:
all regions were either double-stranded or sequenced with an
alternate chemistry or covered by high quality data (i.e., Phred
quality >= 30); an attempt was made to resolve all sequencing
problems, such as compressions and repeats; all regions were
covered by at least one plasmid subclone or more than one M13
subclone; and the assembly was confirmed by restriction digest.

Sequence Validation:

FEATURES
source
1..70560
/organism="Homo sapiens"

Location/Qualifiers
1..70560
/organism="Homo sapiens"

ECORI
BglII
HindIII

SeqDerMap	EngPrnt	SeqDerMap	EngPrnt	SeqDerMap	EngPrnt	SeqDerMap	EngPrnt
2361	2367	13593	13641	9348	9355		
8766	8774	3016	3019	2067	2083		
355	<800	118	<800	5415	5469		
1862	1829	984	1034	292	<800		
3108	3103	3470	3420	1601	1599		
1824	1829	2939	3019	4544	4574		
14428	14564	3419	3420	1417	1428		
418	<800	631	<800	10351	10352		
2391	2367	4397	4299	1357	1341		
944	956	3595	3579	4150	4220		
2372	2367	3583	3579	3095	3123		
342	<800	1163	1141	13617	13613		
580	<800	3327	3420	1117	1135		
1054	1058	1517	1453	4608	4574		
2164	2178	1451	1453	4549	4574		
731	<800	1054	1034	4013	4012		
5580	5583	554	<800	4756	4574		
7525	7519	5987	5947	1298	1275		
1982	1969	7258	7182	1514	1508		
12841	12674	1478	1453				
536	<800	1117	1141				
1407	1384	1364	1366				
40	<800	3095	3019				
1856	1829	728	<800				
180	<800	4140	4059				
2254	2178	1871	1844				
1208	1199	3260	3284				

complement (284469.1:44135. .44337),
complement (284469.1:40468. .40797))
/gene="WNK3"
join(complement (6577. .7113),
complement (AL049793.4:8687. .8859),
complement (AL049793.4:6663. .6883),
complement (AL049793.4:5490. .5647),
complement (284469.1:143915. .144003),
complement (284469.1:140319. .140538),
complement (284469.1:136697. .136991),
complement (284469.1:135278. .135471),
complement (284469.1:135032. .135181),
complement (284469.1:101418. .101482),
complement (284469.1:97884. .98031),
complement (284469.1:96103. .96219),
complement (284469.1:93676. .93831),
complement (284469.1:93543. .93594),
complement (284469.1:92138. .92275),
complement (284469.1:90841. .91778),
complement (284469.1:81012. .81243),
complement (284469.1:80486. .80616),
complement (284469.1:79082. .79695),
complement (284469.1:74953. .75164),
complement (284469.1:44135. .44337),
complement (284469.1:40468. .40797))
/gene="WNK3"
/locus_tag="RP3-390013.2-002"
/product="WNK lysine deficient protein kinase 3"
/note="match: cDNAs: AY082340.1"
join(complement (6577. .7113),
complement (AL049793.4:8687. .8859),
complement (AL049793.4:6663. .6883),
complement (AL049793.4:5490. .5647),
complement (284469.1:143915. .144003),
complement (284469.1:140319. .140538),
complement (284469.1:136697. .136991),
complement (284469.1:135278. .135471),
complement (284469.1:135032. .135181),
complement (284469.1:101418. .101482),
complement (284469.1:97884. .98031),
complement (284469.1:96103. .96219),
complement (284469.1:93676. .93831),
complement (284469.1:93543. .93594),
complement (284469.1:92138. .92275),
complement (284469.1:90841. .91778),
complement (284469.1:81012. .81243),
complement (284469.1:80486. .80616),
complement (284469.1:79082. .79695),
complement (284469.1:74953. .75164),
complement (284469.1:44135. .44337),
complement (284469.1:40468. .40797))
/gene="WNK3"
/locus_tag="RP3-390013.2-002"
/standard_name="OTTHUMP0000023370"
/note="match: proteins: Q8RCX6"
/codon_start=1
/product="WNK lysine deficient protein kinase 3"
/protein_id="CAI40707.1"
/db_xref="GI:57209549"
/db_xref="GOA:Q5JRC2"
/db_xref="InterPro:IPR000719"
/db_xref="InterPro:IPR001245"
/db_xref="InterPro:IPR002290"
/db_xref="InterPro:IPR008271"
/db_xref="UniProt/TrEMBL:Q5JRC2"
/translation="MAYDGDGPASTDESKPDGISFNRPVQAATLTVEARLKEKNS
TFSSAGTEVERKRFIFSKSVEMTDDKVAESSPKDERIKAMNIPRVDKLPNSVLRGQ
EVKYEQCKSTSIKDCFKNEKEEMEEAAKAVATSPSGRFLKFDIELGRGAFKT
VYKGLDTWVEWACWELQDRKLTAEQQQRFKEEAEMLKGLQHPNIVRFYDWSIELK
GKCKIVLTMLTSGTLTKYTKRFKVMKPKVLSWCRQILKGLQFLHTPTPIIHRDL
KCDNIITITGPTGSKVIGDLGLATLMRTSFAKSVIGTPEFWAPMEYEEHYDESVDVYAF

Query Match 87.2%; Score 21.8; DB 8; Length 73084;
Best Local Similarity 92.0%; Pred. No. 16;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAAAAAATCAGGCCAGGCATGG 25
|||||
Db 3519 AAAAAAAAAACAGGCCAGGCATGG 3543
|||||

RESULT 39
AP000280 73687 bp DNA linear PRI 10-AUG-2005
LOCUS Homo sapiens genomic DNA, chromosome 21q22.1, D21S226-AML region,
DEFINITION clone:CM21-S12, complete sequence.
ACCESSION AP000280
VERSION AP000280.3 GI:71891787
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y. and Sakaki,Y.
Homo sapiens 73,686bp genomic DNA of 21q22.1
Published Only in Database (1999)
REFERENCE 2 (bases 1 to 73687)
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y. and Sakaki,Y.
Direct Submission
Submitted (13-MAY-1999) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC),
Kitasato Univ., 1-15-1 Kitasato, Sagamihara, Kanagawa, 228-8555,
Japan [E-mail:hattori@gsc.riken.go.jp,
URL:http://hgp.gsc.riken.go.jp/, Tel:81-42-778-9923,
Fax:81-42-778-9924]
On Aug 5, 2005 this sequence version replaced gi:6016885.
The sequence is a part of the data (ACCESSION No. AP000174 -
AP000194).
The sequencing project is supported by Japan Science Technology
Corporation (JST) and The Institute of Physical and Chemical
Research (RIKEN).
Evidence from one overlapping clone (RP11-630H12), four mRNAs
(BC069689, BC069709, BC069766, BC069779) and four ESTs (H11586,
BX101811, AI888342, BX419841) indicate that there was an A missing
from the chromosome 21 consensus sequence at position 29630 in this
clone. We confirmed that this clone sequence was correct as
previously submitted by reviewing the original trace data, but
since all other evidence suggests that the base is truly missing we
have inserted an additional base in this submission.

FEATURES
source
1. .73687
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosomes="21"
/map="21q22.1"
/clone="CM21-S12"

ORIGIN

Query Match 87.2%; Score 21.8; DB 8; Length 73687;
Best Local Similarity 92.0%; Pred. No. 16;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAAAAAATCAGGCCAGGCATGG 25
|||||
Db 35397 AAAAAAAAAATCAGGCCAGGCATGG 35421
|||||

RESULT 40
AC026374 75385 bp DNA linear PRI 31-JAN-2002
LOCUS


```
misc_feature /note="match: GSS: Em:AQ311353"
634..768
misc_feature /note="match: GSS: Em:AQ317636"
641..888
misc_feature /note="match: GSS: Em:AQ678277"
complement(648..859)
misc_feature /note="match: STS: Em:G17164"
648..813
misc_feature /note="match: STS: Em:RSA108WE1"
complement(648..785)
misc_feature /note="match: GSS: Em:AQ887510"
650..927
misc_feature /note="match: GSS: Em:AQ869729"
651..859
misc_feature /note="match: GSS: Em:AQ754783"
652..877
misc_feature /note="match: GSS: Em:AQ282607"
complement(654..814)
misc_feature /note="match: GSS: Em:AQ118318"
658..935
misc_feature /note="match: GSS: Em:AQ376808"
661..912
misc_feature /note="match: STS: Em:G40890"
complement(664..931)
misc_feature /note="match: GSS: Em:AQ455403"
complement(676..877)
misc_feature /note="match: GSS: Em:AQ933387"
679..898
misc_feature /note="match: STS: Em:G01739"
680..884
misc_feature /note="match: GSS: Em:AQ589116"
complement(680..877)
misc_feature /note="match: GSS: Em:AZ023541"
684..889
misc_feature /note="match: GSS: Em:AQ378981"
687..850
misc_feature /note="match: GSS: Em:AQ538621"
688..893
misc_feature /note="match: GSS: Em:AQ550571"
complement(688..819)
misc_feature /note="match: GSS: Em:AQ337044"
690..908
misc_feature /note="match: GSS: Em:AZ020968"
690..850
misc_feature /note="match: GSS: Em:AQ341026"
695..833
misc_feature /note="match: STS: Em:G33512 Em:RSA045WH9"
complement(708..934)
misc_feature /note="match: GSS: Em:AQ777297"
complement(708..834)
misc_feature /note="match: GSS: Em:AQ064770"
708..831
misc_feature /note="match: GSS: Em:AQ387275"
complement(728..936)
misc_feature /note="match: STS: Em:AQ268772"
complement(745..911)
misc_feature /note="match: GSS: Em:B64703"
complement(746..864)
misc_feature /note="match: GSS: Em:AQ148978"
776..938
misc_feature /note="match: GSS: Em:AQ984702"
789..922
misc_feature /note="match: GSS: Em:AQ198586"
789..911
misc_feature /note="match: GSS: Em:AQ830058"
1025..1133
repeat_region /note="MIR repeat: matches 100..226 of consensus"
1496..1581
repeat_region /note="L2 repeat: matches 2613..2709 of consensus"
1850..2106
repeat_region /note="MLTJL repeat: matches 207..466 of consensus"
complement(join(2349..2435,5293..5409))
gene /gene="dJ336M4.1"

complement(join(<2349..2435,5293..5409))
/gene="dJ336M4.1"
/product="dJ336M4.1 (Miz-1)"
/note="match: cDNA: Em:U14556 Em:U22396 Em:U22396 Em:Y09723
match: ESTs: Em:AA021389 Em:H31841 Em:AI655871 Em:AA624099
Em:AW478691 Em:AA018620 Em:AW408296 Em:AW294814
Em:AW432021 Em:AW481026"
/evidence=not_experimental
2539..2582
repeat_region /note="L2 repeat: matches 2648..2692 of consensus"
2583..2682
repeat_region /note="L2 repeat: matches 2576..2680 of consensus"
2842..3213
repeat_region /note="L2 repeat: matches 1731..2159 of consensus"
3218..3530
repeat_region /note="AluJo repeat: matches 1..312 of consensus"
3568..3783
repeat_region /note="AluJo repeat: matches 1..196 of consensus"
3793..3874
repeat_region /note="L2 repeat: matches 2181..2268 of consensus"
3883..4185
repeat_region /note="AluX repeat: matches 1..302 of consensus"
4490..4709
repeat_region /note="MIR repeat: matches 17..260 of consensus"
4718..5487
misc_feature /note="CpG island"
/evidence=not_experimental
5044..5255
repeat_region /note="106 copies 2 mer cc 58% conserved"
5576..5794
repeat_region /note="MIR repeat: matches 31..262 of consensus"
5925..6236
repeat_region /note="L2 repeat: matches 373..693 of consensus"
6321..6632
repeat_region /note="AluY repeat: matches 1..309 of consensus"
6646..6771
repeat_region /note="FLAM_A repeat: matches 1..133 of consensus"
6808..6930
repeat_region /note="LIPB2 repeat: matches 6032..6155 of consensus"
6931..7240
repeat_region /note="AluY repeat: matches 1..310 of consensus"
7241..7930
repeat_region /note="LIPB2 repeat: matches 5337..6032 of consensus"
7931..8240
repeat_region /note="AluX repeat: matches 1..301 of consensus"
8241..8966

Query Match 87.2%; Score 21.8; DB 8; Length 76241;
Best Local Similarity 92.0%; Pred. No. 16;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAAAAAATCAGCGCAGCATGG 25
||||||| |||||
Db 50887 AAAAAAAAAAGCAGCGCAGCATGG 50863

RESULT 42
AP0000039 100000 bp DNA linear PRI 16-FEB-2005
LOCUS Homo sapiens genomic DNA, chromosome 21q22.1, segment 10/28,
DEFINITION complete sequence.
ACCESSION AP0000039
VERSION AP0000039.1 GI:3132349
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
TITLE Homo sapiens genomic DNA, chromosome 21q
JOURNAL Published Only in Database (1998)
```


ACCESSION	AL109811	
VERSION	AL109811.40	GI:18152762
KEYWORDS	HTG; CpG island; FK506; FLJ41371; FRAP1; MASP2; PMSCSL2; SRM; TARDBP.	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.	
REFERENCE	1 (bases 1 to 112769)	
AUTHORS	Wallis, J.	
TITLE	Direct Submission	

NOTICE:

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate

JOURNAL

Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vega@sanger.ac.uk
 Clone requests: clonerequest@sanger.ac.uk
 On Jan 15, 2002 this sequence version replaced gi:11967852.
 The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:
 Em.: EMBL; Sw.: SWISSPROT; Tr.: TREMBL; Wp.: WORMPEP; Information on the WORMPEP database can be found at
http://www.sanger.ac.uk/projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping Group. Further information can be found at
<http://www.sanger.ac.uk/HGP/Chr1>
 RP4-635E18 is from the library RPCI-4 constructed by the group of Pieter de Jong. For further details see
<http://www.chori.org/bacpac/home.htm>
 VECTOR: pCYPAC2
 ----- Genome Center
 Center: Wellcome Trust Sanger Institute
 Center code: SC
 Web site: <http://www.sanger.ac.uk>
 Contact: vega@sanger.ac.uk

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

Location/Qualifiers

```
1. .112769
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosomes="1"
/map="p36.11-36.31"
/clone="RP4-635E18"
/clone_lib="RPCI-4"
1
/notes="Clone left end: RP4-635E18"
2691. .3057
/notes="Single clone region. Sequence from reads from a short insert library derived from a single pUC clone. Restriction digest data confirm the assembly."
join(8705. .8776,10064. .10313,13192. .13355,15081. .15221,16777. .16837)
/genes="TARDBP"
/locus_tag="RP4-635E18.2-004"
join(8705. .8776,10064. .10313,13192. .13355,15081. .15221,16777. .16837)
/genes="TARDBP"
/locus_tag="RP4-635E18.2-004"
/product="TAR DNA binding protein"
/notes="match: ESTs: Em:BI547516.1 Em:BI550396.1"
join(8990. .9091,10064. .10313,13192. .13355,15081. .15221,16777. .16947,18472. .20391)
/genes="TARDBP"
/locus_tag="RP4-635E18.2-001"
join(8990. .9091,10064. .10313,13192. .13355,15081. .15221,16777. .16947,18472. .20391)
/genes="TARDBP"
/locus_tag="RP4-635E18.2-001"
/product="TAR DNA binding protein"
/notes="match: ESTs: Em:CD703666.1
match: cDNAs: Em:AL050265.1 Em:BC001487.1 Em:BC001487.2
Em:BC012873.1"
join(9003. .9091,10064. .10313,11872. .11994,13192. .13355,15081. .15221)
/genes="TARDBP"
/locus_tag="RP4-635E18.2-005"
join(9003. .9091,10064. .10313,11872. .11994,13192. .13355,15081. .15221)
```

FEATURES

source

misc_feature

misc_feature

gene

mRNA

gene

mRNA

gene

mRNA

gene

mRNA

gene

mRNA

gene

mRNA

CDS

CDS

CDS

```
/gene="TARDBP"
/locus_tag="RP4-635E18.2-005"
/product="TAR DNA binding protein"
/notes="match: ESTs: Em:BM558331.1 Em:BX509845.1"
join(9020. .9091,10064. .10313,12471. .12602,13192. .13355,15081. .15221,16777. .16946)
/genes="TARDBP"
/locus_tag="RP4-635E18.2-006"
join(9020. .9091,10064. .10313,12471. .12602,13192. .13355,15081. .15221,16777. .16946)
/genes="TARDBP"
/product="TAR DNA binding protein"
/notes="match: ESTs: Em:BE901725.1 Em:BG480879.1
Em:BG484362.1"
join(10064. .10313,13192. .13355,15081. .15221,16777. .16947,18472. .18598,19541. .19677,20412. .21367)
/genes="TARDBP"
/locus_tag="RP4-635E18.2-002"
join(10064. .10313,13192. .13355,15081. .15221,16777. .16947,18472. .18598,19541. .19677,20412. .21367)
/genes="TARDBP"
/locus_tag="RP4-635E18.2-002"
/product="TAR DNA binding protein"
/notes="match: cDNAs: Em:AK045495.1 Em:AK076208.1"
join(10064. .10313,13192. .13355,15081. .15221,16777. .16947,18472. .18589,19541. .20531)
/genes="TARDBP"
/locus_tag="RP4-635E18.2-003"
join(10064. .10313,13192. .13355,15081. .15221,16777. .16947,18472. .18589,19541. .20531)
/genes="TARDBP"
/locus_tag="RP4-635E18.2-003"
/product="TAR DNA binding protein"
/notes="match: ESTs: Em:BC02100.1"
join(10076. .10313,13192. .13355,15081. .15221,16777. .16947,18472. .18598,19541. .19596)
/genes="TARDBP"
/locus_tag="RP4-635E18.2-002"
/standard_names="OTHOMP00000002172"
/notes="match: proteins: Tr:AAP92688 Tr:AAP92691 Tr:Q8R0B4"
/codon_start=1
/product="TAR DNA binding protein"
/protein_id="CAI122100.1"
/db_xref="GI:56204105"
/db_xref="GOA:Q5TEQ7"
/db_xref="UniProt/TrEMBL:Q5TEQ7"
/translation="MSEYIRVTEDENDEPIEIPSEDDGTLLSTVTAQPPGACGLRYR
NPVSCMGVRLVEGILHAPDAGWGNLVVYVYPKDKRKMDETASSAVKVKRAVQK
TSDLLVGLPKTKTEQDLKEYFTSTGEVLMQVKDKLTKGHSKGFYRFTVEYQVK
VMSQRHMIDGRWCDCCKLNSKQSDPEPLRSRVFVGRCTEDMTDELREFFSQYGDVM
DVFIKPPFAFAFVTFDQIAQSLCGEDLIKIGISVHISNAEPKHNSNRQLERSGRF
GGNPVHLISNVYGRSTSLKVL"
join(10076. .10313,13192. .13355,15081. .15221,16777. .16947,18472. .18589,19541. .19596)
/genes="TARDBP"
/locus_tag="RP4-635E18.2-003"
/standard_names="OTHOMP00000002173"
/codon_start=1
/product="TAR DNA binding protein"
/protein_id="CAI122099.1"
/db_xref="GI:56204104"
/db_xref="GOA:Q5TEQ8"
/db_xref="UniProt/TrEMBL:Q5TEQ8"
/translation="MSEYIRVTEDENDEPIEIPSEDDGTLLSTVTAQPPGACGLRYR
NPVSCMGVRLVEGILHAPDAGWGNLVVYVYPKDKRKMDETASSAVKVKRAVQK
TSDLLVGLPKTKTEQDLKEYFTSTGEVLMQVKDKLTKGHSKGFYRFTVEYQVK
VMSQRHMIDGRWCDCCKLNSKQSDPEPLRSRVFVGRCTEDMTDELREFFSQYGDVM
DVFIKPPFAFAFVTFDQIAQSLCGEDLIKIGISVHISNAEPKHNSNRQLERSGRF
GVLHLSNVYGRSTSLKVL"
join(10076. .10313,13192. .13355,15081. .15221,16777. .16947,
```


REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 115909)
Birren,B., Linton,L., Nusbaum,C. and Lander,B.
Homo sapiens chromosome 13, clone RP11-474E23
Unpublished
2 (bases 1 to 115909)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Bida,F.,
Boguslavsky,L., Boukagalter,B., Brown,A., Burkett,G.,
Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
Collimore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczký,J.,
Levine,R., Liu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Testaye-S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (10-AUG-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 115909)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Bida,F., Boguslavsky,L.,
Boukagalter,B., Brown,A., Burkett,G., Campopiano,A., Castle,A.,
Choepel,Y., Colangelo,M., Collins,S., Collimore,A., Cooke,P.,
DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Ferreira,P.,
FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M.,
Graham,L., Grand-Pierre,N., Hagos,B., Heaford,A., Horton,D.,
Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A., LaRocque,K.,
Lamazares,R., Landers,T., Lehoczký,J., Levine,R., Liu,C., Liu,G.,
Macdonald,P., Marquis,N., McCarthy,M., McEwan,P., McKernan,K.,
McPheeters,R., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
Morrow,J., Murphy,T., Naylor,J., Norman,C.H., O'Connor,T.,
O'Donnell,P., O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K.,
Pierre,N., Pisani,C., Pollara,V., Raymond,C., Rieback,M., Riley,R.,
Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Severy,P.,
Sougnuez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J.,
Tirrell,A., Travers,M., Trigilio,J., Vassiliev,H., Viel,R., Vo,A.,
Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J.,
Zimmer,A. and Zody,M.
Direct Submission
Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Sep 23, 2000 this sequence version replaced gi:9930750.
All repeats were identified using RepeatMasker:
Smith, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L10224
Center Clone name: 474_E_23
----- Summary Statistics
Sequencing vector: M13; M77815; 97% of reads
Sequencing vector: Plasmid; n/a; 3% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 101354 bases at least Q40

Consensus quality: 109133 bases at least Q30
Consensus quality: 111925 bases at least Q20
Insert size: 121000; agarose-fp
Insert size: 113209; sum-of-contigs
Quality coverage: 5.0 in Q20 bases; agarose-fp
Quality coverage: 5.3 in Q20 b.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 28 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 632: contig of 632 bp in length
* 633 732: gap of 100 bp
* 733 1761: contig of 1029 bp in length
* 1762 1861: gap of 100 bp
* 1862 2965: contig of 1104 bp in length
* 2966 3065: gap of 100 bp
* 3066 4724: contig of 1659 bp in length
* 4725 4824: gap of 100 bp
* 4825 6125: contig of 1301 bp in length
* 6126 6225: gap of 100 bp
* 6226 7471: contig of 1246 bp in length
* 7472 7571: gap of 100 bp
* 7572 8647: contig of 1076 bp in length
* 8648 8747: gap of 100 bp
* 8748 9829: contig of 1082 bp in length
* 9830 9929: gap of 100 bp
* 9930 10967: contig of 1038 bp in length
* 10968 11067: gap of 100 bp
* 11068 12954: contig of 1887 bp in length
* 12955 13054: gap of 100 bp
* 13055 14354: contig of 1300 bp in length
* 14355 14454: gap of 100 bp
* 14455 15459: contig of 1005 bp in length
* 15460 15559: gap of 100 bp
* 15560 16757: contig of 1198 bp in length
* 16758 16857: gap of 100 bp
* 16858 19014: contig of 2157 bp in length
* 19015 19114: gap of 100 bp
* 19115 21051: contig of 1937 bp in length
* 21052 21151: gap of 100 bp
* 21152 23191: contig of 2040 bp in length
* 23192 23291: gap of 100 bp
* 23292 24664: contig of 1373 bp in length
* 24665 24764: gap of 100 bp
* 24765 27243: contig of 2479 bp in length
* 27244 27343: gap of 100 bp
* 27344 30774: contig of 3431 bp in length
* 30775 30874: gap of 100 bp
* 30875 32413: contig of 1539 bp in length
* 32414 32513: gap of 100 bp
* 32514 38423: contig of 5910 bp in length
* 38424 38523: gap of 100 bp
* 38524 44518: contig of 5935 bp in length
* 44519 44618: gap of 100 bp
* 44619 49673: contig of 5055 bp in length
* 49674 49773: gap of 100 bp
* 49774 57087: contig of 7314 bp in length
* 57088 57187: gap of 100 bp
* 57188 64533: contig of 7246 bp in length
* 64534 74157: contig of 9624 bp in length
* 74158 74257: gap of 100 bp
* 74258 89595: contig of 15338 bp in length
* 89596 89695: gap of 100 bp
* 89696 115909: contig of 26214 bp in length.
Location/Qualifiers
1. .115909
/organism="Homo sapiens"
/mol_type="genomic DNA"

FEATURES
source


```

/db xref="taxon:9606"
/chromosome="13"
/map="13"
/clone="RP11-474E23"
/clone_lib="RPC1-11 Human Male BAC"
1. .632
/notes="assembly_fragment"
/clone_end:SP6
vector_side:left"
633. .732
/estimated_length=100
733. .1761
/notes="assembly_fragment"
1762. .1861
/estimated_length=100
1862. .2965
/notes="assembly_fragment"
2966. .3065
/estimated_length=100
3066. .4724
/notes="assembly_fragment"
4725. .4824
/estimated_length=100
4825. .6125
/notes="assembly_fragment"
6126. .6225
/estimated_length=100
6226. .7471
/notes="assembly_fragment"
7472. .7571
/estimated_length=100
7572. .8647
/notes="assembly_fragment"
8648. .8747
/estimated_length=100
8748. .9829
/notes="assembly_fragment"
9830. .9929
/estimated_length=100
9930. .10967
/notes="assembly_fragment"
10968. .11067
/estimated_length=100
11088. .12954
/notes="assembly_fragment"

Query Match      87.2%; Score 21.8; DB 14; Length 115909;
Best Local Similarity 92.0%; Pred. No. 16;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AAAAAAAAAATCAGGCCAGGCATGG 25
Db 52366 AAAAAAAAAATTCAGGCCAGGCATGG 52342

RESULT 50
CNS01DSM      117449 bp DNA linear PRI 26-APR-2001
LOCUS      Human chromosome 14 DNA sequence BAC C-2325P2 of library CalTech-D
DEFINITION      from chromosome 14 of Homo sapiens (Human), complete sequence.
ACCESSION      AL121820
VERSION      AL121820.4 GI:13872724
KEYWORDS      HTG.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
REFERENCE      1 (bases 1 to 117449)
AUTHORS      Heilig, R., Petit, J.-L., Vico, V., Dasilva, C., Robert, C., Wincker, P.,
Brottier, P., Catolico, L., Barbe, V., Pelletier, E., Artiguenave, F.,
Levy, M., Eckenberg, R., Bruls, T., deBerardinis, V., Cruaud, C.,
Gyapay, G., Saurin, W. and Weissenbach, J.

Sequencing of the human chromosome 14
Unpublished
2 (bases 1 to 117449)
Genoscope.
Direct Submission
Submitted (26-APR-2001) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE [E-mail : seqref@genoscope.cns.fr]
- Web : www.genoscope.cns.fr)
On Apr 27, 2001 this sequence version replaced gi:8176917.
-----
Center: Genoscope / Centre National de Sequencage
Center code: GS
Web site: http://www.genoscope.cns.fr/
Contact: SeqRef@genoscope.cns.fr
-----
The following BAC sequence is oriented from the T7 to the SP6 end.
Upstream BAC (overlapping the T7 end) : R-1082A1 (AC=AL13313)
Downstream BAC (overlapping the SP6 end) : R-108B17 -----
Summary Statistics
Assembly program: Phrap; version 2.0
Quality coverage: 7.50x in Q20 bases; sum-of-contigs
-----
Overall quality chart :
Range : bases
0 :
1 - 9 :
10 - 19 :
20 - 29 : 1
30 - 39 : 24
40 - 49 : 1973
50 - 59 : 5247
60 - 69 : 6446
70 - 79 : 14771
80 - 89 : 39167
90 - 99 : 49820
-----
Percentage of bases with a quality value >= 40 : 99 %

FEATURES
source
STS
STS
STS
ORIGIN
Query Match 87.2%; Score 21.8; DB 8; Length 117449;
Best Local Similarity 92.0%; Pred. No. 16;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AAAAAAAAAATCAGGCCAGGCATGG 25
Db 52050 AAAAAAAAAATCCAGGCCAGGCATGG 52074

Search completed: February 3, 2006, 21:25:46
Job time : 877 secs

```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 3, 2006, 10:53:33 ; Search time 203.111 Seconds
(without alignments)
820.326 Million cell updates/sec

Title: US-10-719-900-1

Perfect score: 25
Sequence: 1 aaaaaaaatcacggccagcgatgg 25

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%
Listing first 150 summaries

Database : N Geneseq_21.*

- 1: Geneseqn1980s.*
- 2: Geneseqn1990s.*
- 3: Geneseqn2000s.*
- 4: Geneseqn2001as.*
- 5: Geneseqn2001bs.*
- 6: Geneseqn2002as.*
- 7: Geneseqn2002bs.*
- 8: Geneseqn2003as.*
- 9: Geneseqn2003bs.*
- 10: Geneseqn2003cs.*
- 11: Geneseqn2003ds.*
- 12: Geneseqn2004as.*
- 13: Geneseqn2004bs.*
- 14: Geneseqn2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	21.8	87.2	341	AAK74376	Aak74376 Human imm
2	21.8	87.2	464	ADD35001	Add35001 Mouse mit
3	21.8	87.2	2825	ADL06341	Adl06341 cDNA enco
4	21.8	87.2	3067	AAH19210	Aah19210 Human sec
5	21.8	87.2	3067	ADD67600	Add67600 Human Ly1
6	21.8	87.2	3105	AAH19178	Aah19178 Human sec
7	21.8	87.2	3288	ACC72030	Acc72030 BC00571 g
8	21.8	87.2	5017	AAK69927	Aak69927 Human imm
9	21.8	87.2	5017	AAK69928	Aak69928 Human imm
10	21.8	87.2	98642	ACN44584	Acn44584 Mouse gen
11	21.8	87.2	110000	ACN45090_0	Acn45090 Human gen
12	20.8	83.2	25574	AAK79671	Aak79671 Human imm
13	20.8	83.2	25574	AAK83760	Aak83760 Human imm
14	20.8	83.2	25574	AAK05619	Aal05619 Human rep
15	20.8	83.2	25576	AAK85305	Aak85305 Human imm
16	20.8	83.2	25576	AAK79669	Aak79669 Human imm
17	20.8	83.2	25576	AAK83758	Aak83758 Human imm
18	20.8	83.2	25576	AAK05618	Aal05618 Human rep
19	20.8	83.2	110000	ACN44934_1	Continuation (2 of

Continuation (4 of	14	ADZ46976_3	83.2	110000	20.8	20.2	80.8	12312	5	AAF97872
Adl18927 Human sof	12	ADQ18927	81.6	210204	20.4	20.2	80.8	10236	4	AAQ18927
Aac17433 Human sec	3	AAC17433	80.8	138	20.2	20.2	80.8	10236	4	AAQ17433
Aac14155 Human sec	206	AAC14155	80.8	206	20.2	20.2	80.8	10236	4	AAQ14155
Ac156670 Human col	302	ACL56670	80.8	302	20.2	20.2	80.8	10236	4	AAQ156670
Aak74375 Human imm	339	AAK74375	80.8	339	20.2	20.2	80.8	10236	4	AAQ74375
Aas37660 Novel hum	344	AAS37660	80.8	344	20.2	20.2	80.8	10236	4	AAQ37660
Abk95389 Human ret	384	ABK95389	80.8	384	20.2	20.2	80.8	10236	4	AAQ95389
Aak83882 Human imm	392	AAK83882	80.8	392	20.2	20.2	80.8	10236	4	AAQ83882
Aai187906 Human pol	403	AAI87906	80.8	403	20.2	20.2	80.8	10236	4	AAQ87906
Ab182000 Human ova	465	ABL82000	80.8	465	20.2	20.2	80.8	10236	4	AAQ82000
Acf85697 Human SIR	465	ACF85697	80.8	465	20.2	20.2	80.8	10236	4	AAQ85697
Abn61750 Human can	480	ABN61750	80.8	480	20.2	20.2	80.8	10236	4	AAQ61750
Ac160088 Human col	480	ACL60088	80.8	480	20.2	20.2	80.8	10236	4	AAQ60088
Abas59354 Human foe	484	ABA59354	80.8	484	20.2	20.2	80.8	10236	4	AAQ59354
Aai39155 Probe #78	484	AAI39155	80.8	484	20.2	20.2	80.8	10236	4	AAQ39155
Aak33368 Human bon	484	AAK33368	80.8	484	20.2	20.2	80.8	10236	4	AAQ33368
Aak07570 Human bra	484	AAK07570	80.8	484	20.2	20.2	80.8	10236	4	AAQ07570
Abas33124 Human liv	484	ABS33124	80.8	484	20.2	20.2	80.8	10236	4	AAQ33124
Abas08206 Human gen	526	ABS08206	80.8	526	20.2	20.2	80.8	10236	4	AAQ08206
Abn64674 Human can	666	ABN64674	80.8	666	20.2	20.2	80.8	10236	4	AAQ64674
Aaa35120 Human ade	666	AAA35120	80.8	666	20.2	20.2	80.8	10236	4	AAQ35120
Aaf21242 Human low	666	AAF21242	80.8	666	20.2	20.2	80.8	10236	4	AAQ21242
Abz66936 Human nuc	666	ABZ66936	80.8	666	20.2	20.2	80.8	10236	4	AAQ66936
Abd20785 Human pul	666	ABD20785	80.8	666	20.2	20.2	80.8	10236	4	AAQ20785
Aas28505 Genomic s	1131	AAS28505	80.8	1131	20.2	20.2	80.8	10236	4	AAQ28505
Aas29996 Human lun	1131	AAS29996	80.8	1131	20.2	20.2	80.8	10236	4	AAQ29996
Adg33333 Human nov	1131	ADB33333	80.8	1131	20.2	20.2	80.8	10236	4	AAQ33333
Adg41701 Human res	1131	ADG41701	80.8	1131	20.2	20.2	80.8	10236	4	AAQ41701
Adi97475 Human res	1131	ADI97475	80.8	1131	20.2	20.2	80.8	10236	4	AAQ97475
Abq75320 Human lun	1264	ABQ75320	80.8	1264	20.2	20.2	80.8	10236	4	AAQ75320
Abz58228 Human G-p	1271	ABZ58228	80.8	1271	20.2	20.2	80.8	10236	4	AAQ58228
Acn37887 Tumour-as	1935	ACN37887	80.8	1935	20.2	20.2	80.8	10236	4	AAQ37887
Aak85252 Human imm	2713	AAK85252	80.8	2713	20.2	20.2	80.8	10236	4	AAQ85252
Abz74454 Secreted	3310	ABZ74454	80.8	3310	20.2	20.2	80.8	10236	4	AAQ74454
Ada98878 Human sec	3310	ADA98878	80.8	3310	20.2	20.2	80.8	10236	4	AAQ98878
Adc20907 Human sec	3310	ADC20907	80.8	3310	20.2	20.2	80.8	10236	4	AAQ20907
Abt16993 Human sec	3310	ABT16993	80.8	3310	20.2	20.2	80.8	10236	4	AAQ16993
Abz68001 Human sec	3310	ABZ68001	80.8	3310	20.2	20.2	80.8	10236	4	AAQ68001
Adg46497 Novel hum	3706	ADG46497	80.8	3706	20.2	20.2	80.8	10236	4	AAQ46497
Abq81729 Benign pr	3943	ABQ81729	80.8	3943	20.2	20.2	80.8	10236	4	AAQ81729
Aak68715 Human imm	4013	AAK68715	80.8	4013	20.2	20.2	80.8	10236	4	AAQ68715
Aav99094 DNA methy	4459	AAV99094	80.8	4459	20.2	20.2	80.8	10236	4	AAQ99094
Adg63025 Novel hum	5003	ADG63025	80.8	5003	20.2	20.2	80.8	10236	4	AAQ63025
Aas28641 Genomic s	6186	AAS28641	80.8	6186	20.2	20.2	80.8	10236	4	AAQ28641
Adg41837 Human res	6186	ADG41837	80.8	6186	20.2	20.2	80.8	10236	4	AAQ41837
Adi97611 Human res	6186	ADI97611	80.8	6186	20.2	20.2	80.8	10236	4	AAQ97611
Aas28642 Genomic s	6191	AAS28642	80.8	6191	20.2	20.2	80.8	10236	4	AAQ28642
Aas28643 Genomic s	6191	AAS28643	80.8	6191	20.2	20.2	80.8	10236	4	AAQ28643
Adg41838 Human res	6191	ADG41838	80.8	6191	20.2	20.2	80.8	10236	4	AAQ41838
Adg41839 Human res	6191	ADG41839	80.8	6191	20.2	20.2	80.8	10236	4	AAQ41839
Adi97613 Human res	6191	ADI97613	80.8	6191	20.2	20.2	80.8	10236	4	AAQ97613
Adi97612 Human res	6191	ADI97612	80.8	6191	20.2	20.2	80.8	10236	4	AAQ97612
Aal44086 Human tra	6339	AAL44086	80.8	6339	20.2	20.2	80.8	10236	4	AAQ44086
Aak70928 Human imm	6743	AAK70928	80.8	6743	20.2	20.2	80.8	10236	4	AAQ70928
Aal05842 Human rep	6743	AAL05842	80.8	6743	20.2	20.2	80.8	10236	4	AAQ05842
Ab198406 Human tes	6743	ABL98406	80.8	6743	20.2	20.2	80.8	10236	4	AAQ98406
Abk46568 DNA enco	7426	ABK46568	80.8	7426	20.2	20.2	80.8	10236	4	AAQ46568
Aak84896 Human imm	7794	AAK84896	80.8	7794	20.2	20.2	80.8	10236	4	AAQ84896
Aas28608 Genomic s	7794	AAS28608	80.8	7794	20.2	20.2	80.8	10236	4	AAQ28608
Adg41804 Human res	7794	ADG41804	80.8	7794	20.2	20.2	80.8	10236	4	AAQ41804
Adi97578 Human res	7794	ADI97578	80.8	7794	20.2	20.2	80.8	10236	4	AAQ97578
Aal51033 Human Ras	9454	AAL51033	80.8	9454	20.2	20.2	80.8	10236	4	AAQ51033
Aas28920 Human imm	10236	AAS28920	80.8	10236	20.2	20.2	80.8	10236	4	AAQ28920
Aas28880 Human imm	10236	AAS28880	80.8	10236	20.2	20.2	80.8	10236	4	AAQ28880
Aak68907 Human imm	10236	AAK68907	80.8	10236	20.2	20.2	80.8	10236	4	AAQ68907
Aak80678 Human imm	10236	AAK80678	80.8	10236	20.2	20.2	80.8	10236	4	AAQ80678
Aak68717 Human imm	10236	AAK68717	80.8	10236	20.2	20.2	80.8	10236	4	AAQ68717
Abas06792 Human gen	10236	ABA06792	80.8	10236	20.2	20.2	80.8	10236	4	AAQ06792
Abv84129 Human pol	10236	ABV84129	80.8	10236	20.2	20.2	80.8	10236	4	AAQ84129
Adb31721 Human nov	10236	ADB31721	80.8	10236	20.2	20.2	80.8	10236	4	AAQ31721
Adb31761 Human nov	10236	ADB31761	80.8	10236	20.2	20.2	80.8	10236	4	AAQ31761
Aaf97872 Human neu	12312	AAF97872	80.8	12312	20.2	20.2	80.8	10236	4	AAQ97872

```
C 93 20.2 80.8 13224 4 AAS41751
C 94 20.2 80.8 13224 4 ABA06811
C 95 20.2 80.8 13224 6 ABV84148
C 96 20.2 80.8 15000 6 ABL62905
C 97 20.2 80.8 15866 4 AAK90892
C 98 20.2 80.8 16500 12 ADP68405
C 99 20.2 80.8 17425 2 AAX28366
C 100 20.2 80.8 19301 3 AAX35123
C 101 20.2 80.8 19301 3 AAF21245
C 102 20.2 80.8 19301 10 ABZ96939
C 103 20.2 80.8 19301 11 ABD20788
C 104 20.2 80.8 19440 2 AAV99129
C 105 20.2 80.8 19802 4 AAK90894
C 106 20.2 80.8 19802 4 AAL04554
C 107 20.2 80.8 21403 4 ABL97477
C 108 20.2 80.8 21403 4 AAK90893
C 109 20.2 80.8 21403 4 AAL04555
C 110 20.2 80.8 21403 4 ABL97478
C 111 20.2 80.8 23763 4 AAK90895
C 112 20.2 80.8 23763 4 AAL04558
C 113 20.2 80.8 23763 4 ABL97481
C 114 20.2 80.8 24977 4 AAL03263
C 115 20.2 80.8 24977 4 AAL05788
C 116 20.2 80.8 24977 4 ABL98352
C 117 20.2 80.8 24983 4 AAL05787
C 118 20.2 80.8 24983 4 AAL03262
C 119 20.2 80.8 24983 4 ABL98351
C 120 20.2 80.8 26354 12 ADQ97798
C 121 20.2 80.8 27062 4 AAK71622
C 122 20.2 80.8 27062 4 AAL05346
C 123 20.2 80.8 27062 4 ABL98215
C 124 20.2 80.8 27990 11 ACN44366
C 125 20.2 80.8 28449 12 ADQ97240
C 126 20.2 80.8 29329 4 AAK78512
C 127 20.2 80.8 29329 4 AAK70791
C 128 20.2 80.8 29329 5 ABA18026
C 129 20.2 80.8 29329 5 ABA20511
C 130 20.2 80.8 30573 4 AAL04181
C 131 20.2 80.8 32167 4 AAL05509
C 132 20.2 80.8 32191 4 AAL199472
C 133 20.2 80.8 32191 4 AAI64117
C 134 20.2 80.8 32191 12 ADF71715
C 135 20.2 80.8 32195 4 AAL04180
C 136 20.2 80.8 37262 11 ACN43890
C 137 20.2 80.8 43256 12 ADQ97736
C 138 20.2 80.8 44840 4 AAK71803
C 139 20.2 80.8 44978 13 ADS36458
C 140 20.2 80.8 45546 2 AAX23520
C 141 20.2 80.8 47243 11 ACN44690
C 142 20.2 80.8 52616 4 AAK70459
C 143 20.2 80.8 52616 4 AAK78930
C 144 20.2 80.8 52710 13 ABD33436
C 145 20.2 80.8 57038 11 ADP65343
C 146 20.2 80.8 57130 8 ABO77243
C 147 20.2 80.8 65559 11 ACN44408
C 148 20.2 80.8 65608 6 ABL62910
C 149 20.2 80.8 65608 6 ABL64414
C 150 20.2 80.8 65608 6 ABL67668
```

ALIGNMENTS

```
RESULT 1
ID AAK74376
XX AAK74376 standard; DNA; 341 BP.
AC AAK74376;
XX
DT 07-NOV-2001 (first entry)
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:29188.
XX
```

```
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX Homo sapiens.
OS
PN WO200157182-A2.
XX
XX 09-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US001354.
XX
XX 31-JAN-2000; 2000US-0179065P.
XX 04-FEB-2000; 2000US-0180628P.
XX 24-FEB-2000; 2000US-0184664P.
XX 02-MAR-2000; 2000US-0186350P.
XX 16-MAR-2000; 2000US-0189874P.
XX 17-MAR-2000; 2000US-0190076P.
XX 18-APR-2000; 2000US-0198123P.
XX 19-MAY-2000; 2000US-0205515P.
XX 07-JUN-2000; 2000US-0209467P.
XX 28-JUN-2000; 2000US-0214886P.
XX 30-JUN-2000; 2000US-0215135P.
XX 07-JUL-2000; 2000US-0216647P.
XX 07-JUL-2000; 2000US-0216880P.
XX 11-JUL-2000; 2000US-0217487P.
XX 11-JUL-2000; 2000US-0217496P.
XX 14-JUL-2000; 2000US-0218290P.
XX 26-JUL-2000; 2000US-0220963P.
XX 26-JUL-2000; 2000US-0220964P.
XX 14-AUG-2000; 2000US-0224518P.
XX 14-AUG-2000; 2000US-0224519P.
XX 14-AUG-2000; 2000US-0225213P.
XX 14-AUG-2000; 2000US-0225214P.
XX 14-AUG-2000; 2000US-0225266P.
XX 14-AUG-2000; 2000US-0225267P.
XX 14-AUG-2000; 2000US-0225268P.
XX 14-AUG-2000; 2000US-0225270P.
XX 14-AUG-2000; 2000US-0225447P.
XX 14-AUG-2000; 2000US-0225757P.
XX 14-AUG-2000; 2000US-0225758P.
XX 14-AUG-2000; 2000US-0225759P.
XX 18-AUG-2000; 2000US-0226279P.
XX 22-AUG-2000; 2000US-0226681P.
XX 22-AUG-2000; 2000US-0226868P.
XX 22-AUG-2000; 2000US-0227182P.
XX 23-AUG-2000; 2000US-0227009P.
XX 30-AUG-2000; 2000US-0228924P.
XX 01-SEP-2000; 2000US-0229287P.
XX 01-SEP-2000; 2000US-0229343P.
XX 01-SEP-2000; 2000US-0229344P.
XX 01-SEP-2000; 2000US-0229345P.
XX 05-SEP-2000; 2000US-0229509P.
XX 05-SEP-2000; 2000US-0229513P.
XX 06-SEP-2000; 2000US-0230437P.
XX 06-SEP-2000; 2000US-0230438P.
XX 08-SEP-2000; 2000US-0231242P.
XX 08-SEP-2000; 2000US-0231243P.
XX 08-SEP-2000; 2000US-0231244P.
XX 08-SEP-2000; 2000US-0231413P.
XX 08-SEP-2000; 2000US-0231414P.
XX 08-SEP-2000; 2000US-0232080P.
XX 08-SEP-2000; 2000US-0232081P.
XX 12-SEP-2000; 2000US-0231968P.
XX 14-SEP-2000; 2000US-0232397P.
XX 14-SEP-2000; 2000US-0232398P.
XX 14-SEP-2000; 2000US-0232399P.
XX 14-SEP-2000; 2000US-0232400P.
XX 14-SEP-2000; 2000US-0232401P.
XX 14-SEP-2000; 2000US-0233063P.
XX 14-SEP-2000; 2000US-0233064P.
XX 14-SEP-2000; 2000US-0233065P.
XX 21-SEP-2000; 2000US-0234223P.
XX 21-SEP-2000; 2000US-0234274P.
```

```
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0234984P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0246177P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.

PR 05-JAN-2001; 2001US-0259678P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-483426/52.
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and metastasis.
XX Disclosure; SEQ ID NO 29188; 3071pp + Sequence Listing; English.
XX AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting the
CC nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
CC represent sequences used in the exemplification of the present invention
XX SQ Sequence 341 BP; 133 A; 49 C; 77 G; 82 T; 0 U; 0 Other;

Query Match 87.2%; Score 21.8; DB 4; Length 341;
Best Local Similarity 92.0%; Pred. No. 7.6;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAAAAAATCAGCGCCAGGCATGG 25
Db 236 AAAAAAAAAAACAGGCCAGGCATGG 260

RESULT 2
ADD35001
ID ADD35001 standard; DNA; 464 BP.
XX AC ADD35001;
XX 15-JAN-2004 (first entry)
XX Mouse mitochondrial DNA sequence SEQ ID NO:2781.
XX db; mouse; array; mitochondrial; hybridisation; energy-metabolism;
XX mitochondrial disease; oxidative phosphorylation dysfunction;
XX oxidative stress; apoptosis; aging.
XX Mus musculus.
XX WO2003020220-A2.
XX 13-MAR-2003.
XX 30-AUG-2002; 2002WO-US027886.
XX 30-AUG-2001; 2001US-0316323P.
XX 31-AUG-2001; 2001CA-02356540.
XX (UYEM-) UNIV EMORY.
XX Wallace DC, Levy S, Kerstann K, Procaccio V;
XX WPI; 2003-300821/29.
```

PT Array containing probes for genes involved in mitochondrial biology,
 PT useful for determining mitochondrial biology gene expression profiles for
 PT use in diagnosing pathologies and identifying biochemical pathways.

XX
 XX Claim 2; SEQ ID NO 2781; 201pp; English.

CC The invention relates to a novel array comprising at least two isolated
 CC nucleotide molecules, each molecule having a sequence capable of uniquely
 CC hybridising to a nucleic acid molecule which is an expression product of
 CC a gene involved in mitochondrial biology. The array comprises two or more
 CC isolated nucleic acid molecules or spots, each molecule having a sequence
 CC chosen from sequence of 994 human probes and 2046 mouse probes. An array
 CC of the invention is useful for determining an expression profile of a
 CC mouse or human sample containing nucleic acid, by contacting the array
 CC with the sample under conditions allowing selective hybridisation, and
 CC measuring hybridisation of nucleic acid in the sample to the array to
 CC produce an expression profile. The array is also useful for determining
 CC an expression profile of a first labelled sample containing nucleic acid
 CC relative to a second, differently labelled sample containing nucleic
 CC acid. The second sample is a reference or a standard. An array is useful
 CC for determining an expression profile diagnostic of an energy-metabolism-
 CC related physiological condition. An array of the invention is useful for
 CC determining mitochondrial biology gene expression profiles of organisms,
 CC such as human, mice and closely related species, tissue and organs of
 CC such organisms, which are useful for determining expression profiles
 CC diagnostic of energy metabolism-related physiological conditions.
 CC diagnosing such physiological conditions, identifying biochemical
 CC pathways, genes, and mutations involved in such physiological conditions,
 CC identifying therapeutic agents useful for preventing and/or treating such
 CC physiological conditions, evaluating and/or monitoring the efficacy of
 CC such therapies, and creating and identifying animal models of human
 CC energy metabolism-related physiological conditions. An array is also
 CC useful for defining expression signatures or profiles for mitochondrial
 CC diseases, as well as distinguishing clinical disorders that result from
 CC oxidative phosphorylation (OXPHOS) dysfunction, oxidative stress,
 CC apoptosis and aging. An array of the invention contains probes of genes
 CC not previously recognised to participate in mitochondrial biology. The
 CC sequences shown in ADP33224-ADP35260 represent murine mitochondrial DNA
 CC clones used to make the probes of the invention. Some sequences are not
 CC present, these are SEQ ID NO's 295, 1174, 1213, 1700, 1728, 1730, 1905,
 CC 1906, 2408 and 2643.

XX
 XX Sequence 464 BP; 180 A; 91 C; 80 G; 111 T; 0 U; 2 Other;

Query Match 87.2%; Score 21.8; DB 10; Length 464;
 Best Local Similarity 92.0%; Pred. No. 8;
 Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAAAAAATCAGCGCCAGGCATGG 25

Db 255 AAAAAAAAAATCTTGGCCAGGCATGG 279

RESULT 3

ADL06341/c

ID ADL06341 standard; cDNA; 2825 BP.

AC ADL06341;

XX 06-MAY-2004 (first entry)

XX cDNA encoding human protein-13.2, a site-specific recombinase.

XX Human; protein-13.2; site-specific recombinase;
 KW growth development disorder; tumour; gene; ss.

XX Homo sapiens.

XX Key Location/Qualifiers
 FT 1016..1378
 CDS /*tag= a

FT /product= "Protein-13.2"

XX

PN CN1393548-A.

XX 29-JAN-2003.

XX 29-JUN-2001; 2001CN-001131178.

XX 29-JUN-2001; 2001CN-001131178.

XX (BIOW-) BIOWINDOW GENE DEV INC SHANGHAI.

XX Mao Y, Xie Y;

XX WPI; 2003-422181/40.

XX P-PSDB; ADL06342.

XX Polypeptide-human protein-13.2 containing site specific recombinase

XX characteristic sequence fragment and polynucleotide for coding it.

XX Claim 6; SEQ ID NO 1; 32pp; Chinese.

XX The present invention relates to the isolation of human protein-13.2

XX containing a site-specific recombinase characteristic sequence fragment,

XX and the polynucleotide sequence encoding it. Also disclosed is a process

XX for preparing the polypeptide by a DNA recombination technique and

XX application of the polypeptide and polynucleotide in treating diseases

XX such as growth development disorders and tumours. The present sequence

XX encodes human protein-13.2.

XX Sequence 2825 BP; 799 A; 556 C; 707 G; 763 T; 0 U; 0 Other;

XX Query Match 87.2%; Score 21.8; DB 10; Length 2825;

XX Best Local Similarity 92.0%; Pred. No. 11;

XX Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAAAAAATCAGCGCCAGGCATGG 25

Db 291 AAAAAAAAAATTAAGCCAGGCATGG 267

RESULT 4

AAH19210/c

ID AAH19210 standard; cDNA; 3067 BP.

XX AC AAH19210;

XX 17-JUL-2001 (first entry)

XX Human secreted protein-encoding gene 9 cDNA clone HTOHB55, SEQ ID NO:51.

XX Human; secreted protein; proliferative disorder; cancer;

XX foetal abnormality; developmental abnormality; haematopoietic disorder;

XX immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;

XX inflammation; allergy; neurological disorder; Alzheimer's disease;

XX Parkinson's disease; cognitive disorder; schizophrenia; asthma;

XX skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;

XX cardiovascular disorder; angioenic disorder; kidney disorder;

XX gastrointestinal disorder; pregnancy-related disorder; tumour;

XX endocrine disorder; infection; wound healing; vulnerability; cell culture;

XX chemotaxis; food additive; binding partner identification; ss.

XX Homo sapiens.

XX WO200132910-A2.

XX 10-MAY-2001.

XX 25-OCT-2000; 2000WO-US029362.

XX 29-OCT-1999; 99US-0162240P.

XX 30-JUN-2000; 2000US-0215131P.

XX (HUMA-) HUMAN GENOME SCI INC.

PI Ruben SM, Komatsoulis GA, Birse CE, Ni J, Soppet DR;
XX WPI; 2001-335835/35.
DR N-PSDB; AAG62196.
XX
XX Novel 27 isolated human secreted proteins and polynucleotides encoding
PT them useful for treating, diagnosing, preventing Alzheimer's disease,
PT Parkinson's disease, AIDS, rheumatoid arthritis, asthma.
XX
XX Claim 4; Page 469-470; 594pp; English.
XX
XX AAH19170-AAH19250 represent cDNAs corresponding to 27 human secreted
CC protein genes, and AAG62156-AAG62235 represent the proteins they encode.
CC AG62236-AAG62293 represent human secreted protein fragments. The genes
CC and their corresponding secreted proteins are useful for preventing,
CC treating or ameliorating medical conditions, e.g., by protein or gene
CC therapy. Pathological conditions can be diagnosed by determining the
CC amount of the new protein in a sample or by determining the presence of
CC mutations in the new genes. Specific uses are described for each of the
CC 52 genes, based on the tissues in which they are most highly expressed,
CC and include developing products for the diagnosis or treatment of
CC proliferative disorders, cancer, tumours, foetal and developmental
CC abnormalities, haematopoietic disorders, diseases of the immune system,
CC AIDS, autoimmune diseases (e.g., rheumatoid arthritis), inflammation,
CC allergies, neurological disorders (e.g., Alzheimer's disease,
CC Parkinson's disease), cognitive disorders, schizophrenia, asthma, skin
CC disorders (e.g., psoriasis), sepsis, diabetes, atherosclerosis,
CC cardiovascular disorders, angiotensin-related disorders, kidney disorders,
CC gastrointestinal disorders, pregnancy-related disorders, endocrine
CC disorders, and infections. The proteins can also be used to aid wound
CC healing and epithelial cell proliferation, to prevent skin aging due to
CC sunburn, to maintain organs before transplantation, for supporting cell
CC culture of primary tissues, to regenerate tissues, to identify their
CC cognate ligands or binding partners, and in chemotaxis, and can be used
CC as a food additive or preservative to modify storage properties.
CC Antibodies specific for a protein of the invention can be used in
CC alleviating symptoms associated with the disorders mentioned above, and
CC in diagnostic immunoassays e.g., radioimmunoassay or enzyme linked
CC immunosorbent assay (ELISA). The present sequence represents a human
CC secreted protein-encoding cDNA of the invention
XX
SQ Sequence 3067 BP; 672 A; 826 C; 783 G; 786 T; 0 U; 0 Other;

Query Match 87.2%; Score 21.8; DB 4; Length 3067;
Best Local Similarity 92.0%; Pred. No. 11;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAAAAAATCAGCGCCAGGCATGG 25
Db 1184 AAAAAATAATCAGCGCCAGGCACGG 1160

RESULT 5
ADD67600/c
ID ADD67600 standard; cDNA; 3067 BP.
XX
AC ADD67600;
XX
XX 15-JAN-2004 (first entry)
DT
XX
DE Human Ly1485P partial cDNA SEQ ID NO:77.
XX
XX haematological malignancy; immunoconjugate; cytostatic; immunostimulant;
KW vaccine; immunotherapy; cancer; multiple myeloma cell;
KW chronic lymphocytic leukaemia; B cell leukaemia; lymphoma; anti-cancer;
KW human; gene; ss.
XX
XX Homo sapiens.
OS
XX WO2003062401-A2.
PN
XX 31-JUL-2003.
PD
XX

PF 22-JAN-2003; 2003WO-US002353.
XX
PR 22-JAN-2002; 2002US-00057475.
XX
XX (CORI-) CORIXA CORP.
PA
XX Gaiger A, Algate PA, Mannion J, Clapper JD, Wang A, Ordenez N;
PI Carter L, McNeill PD;
XX
XX WPI; 2003-598749/56.
DR
XX
XX New hematological malignancy-related genes and polypeptides, useful for
PT screening anti-cancer agents, and generating antibodies or
PT immunoconjugates for treating e.g. multiple myeloma cell or chronic
PT lymphocytic leukemia.
XX
XX Claim 1; SEQ ID NO 77; 307pp; English.
XX
XX The present invention describes an isolated polynucleotide (I), which is
CC overexpressed in haematological malignancies, and which encodes a
CC polypeptide or an immunogenic fragment of the polypeptide. Also
CC described: (1) an isolated polypeptide; (2) an expression vector
CC comprising (i) operably linked to an expression control sequence; (3) a
CC host cell comprising an expression vector; (4) an isolated antibody that
CC specifically binds to the polypeptide or its immunogenic fragment; and
CC (5) immunoconjugates comprising the antibody above, or an antibody that
CC specifically binds to a polypeptide, or its immunogenic fragment, encoded
CC by (1). (I) has cytostatic and immunostimulant activities, and can be
CC used in vaccines and immunotherapy. The immunoconjugates are useful in
CC the manufacture of a medicament, particularly as active ingredients in a
CC composition for treating cancer, e.g. multiple myeloma cell, chronic
CC lymphocytic leukaemia, B cell leukaemias, or lymphomas in humans, sheep,
CC primates, goats, bovines, equines, porcines, lupines, canines or felines.
CC The polynucleotide (I) or polypeptide can be used for screening anti-
CC cancer agents, and generating antibodies or immunoconjugates for treating
CC or preventing the above-mentioned diseases. The polynucleotide,
CC polypeptide or antibody can be used for detecting, diagnosing or
CC prognosticating the haematological malignancies described above. The
CC present sequence is used in the exemplification of the present invention.
XX
SQ Sequence 3067 BP; 672 A; 826 C; 783 G; 786 T; 0 U; 0 Other;

Query Match 87.2%; Score 21.8; DB 10; Length 3067;
Best Local Similarity 92.0%; Pred. No. 11;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAAAAAATCAGCGCCAGGCATGG 25
Db 1184 AAAAAATAATCAGCGCCAGGCACGG 1160

RESULT 6
AAH19178/c
ID AAH19178 standard; cDNA; 3105 BP.
XX
AC AAH19178;
XX
XX 17-JUL-2001 (first entry)
DT
XX
DE Human secreted protein-encoding gene 9 cDNA clone HTOHB55, SEQ ID NO:19.
XX
XX Human secreted protein; proliferative disorder; cancer;
KW foetal abnormality; developmental abnormality; haematopoietic disorder;
KW immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
KW inflammation; allergy; neurological disorder; Alzheimer's disease;
KW Parkinson's disease; cognitive disorder; schizophrenia; asthma;
KW skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;
KW cardiovascular disorder; angiotensin-related disorder; kidney disorder;
KW gastrointestinal disorder; pregnancy-related disorder; tumour;
KW endocrine disorder; infection; wound healing; vulnerability; cell culture;
KW chemotaxis; food additive; binding partner identification; ss.
XX
XX Homo sapiens.
OS

```
XX WO200132910-A2.
XX 10-MAY-2001.
XX
XX 25-OCT-2000; 2000WO-US029362.
XX
XX 29-OCT-1999; 99US-0162240P.
XX 30-JUN-2000; 2000US-0215131P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Ruben SM, Komatsoulis GA, Birse CE, Ni J, Soppet DR;
XX WPI; 2001-335835/35.
XX N-PSDB; AAG62164.
XX
XX Novel 27 isolated human secreted proteins and polynucleotides encoding
XX PT them useful for treating, diagnosing, preventing Alzheimer's disease,
XX PT Parkinson's disease, AIDS, rheumatoid arthritis, asthma.
XX
XX Claim 4; Page 449-450; 594pp; English.
XX
XX AAH19170-AAH19250 represent cDNAs corresponding to 27 human secreted
XX CC protein genes, and AAG62156-AAG62235 represent the proteins they encode.
XX CC AAG62236-AAG62293 represent human secreted protein fragments. The genes
XX CC and their corresponding secreted proteins are useful for preventing,
XX CC treating or ameliorating medical conditions, e.g., by protein or gene
XX CC therapy. Pathological conditions can be diagnosed by determining the
XX CC amount of the new protein in a sample or by determining the presence of
XX CC mutations in the new genes. Specific uses are described for each of the
XX CC 52 genes, based on the tissues in which they are most highly expressed,
XX CC and include developing products for the diagnosis or treatment of
XX CC proliferative disorders, cancer, tumours, foetal and developmental
XX CC abnormalities, haematopoietic disorders, diseases of the immune system,
XX CC AIDS, autoimmune diseases (e.g., rheumatoid arthritis), inflammation,
XX CC allergies, neurological disorders (e.g., Alzheimer's disease,
XX CC Parkinson's disease), cognitive disorders, schizophrenia, asthma, skin
XX CC disorders (e.g., psoriasis), sepsis, diabetes, atherosclerosis,
XX CC cardiovascular disorders, angiogenic disorders, kidney disorders,
XX CC gastrointestinal disorders, pregnancy-related disorders, endocrine
XX CC disorders, and infections. The proteins can also be used to aid wound
XX CC healing and epithelial cell proliferation, to prevent skin aging due to
XX CC burn, to maintain organs before transplantation, for supporting cell
XX CC culture of primary tissues, to regenerate tissues, to identify their
XX CC cognate ligands or binding partners, and in chemotaxis, and can be used
XX CC as a food additive or preservative to modify storage properties.
XX CC Antibodies specific for a protein of the invention can be used in
XX CC alleviating symptoms associated with the disorders mentioned above, and
XX CC in diagnostic immunoassays e.g., radioimmunoassay or enzyme linked
XX CC immunosorbent assay (ELISA). The present sequence represents a human
XX CC secreted protein-encoding cDNA of the invention
XX
XX SQ Sequence 3105 BP; 678 A; 833 C; 795 G; 791 T; 0 U; 8 Other;
XX
XX Query Match 87.2%; Score 21.8; DB 4; Length 3105;
XX Best Local Similarity 92.0%; Pred. No. 11;
XX Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
XX QY 1 AAAAAAAAAATCAGGCGCAGGCATGG 25
XX ||||| ||||| ||||| ||||| |||||
XX Db 1177 AAAAAAAAAATCAGGCGCAGGCAGG 1153
XX
XX RESULT 7
XX ACC72030
XX ID ACC72030 standard; DNA; 3288 BP.
XX
XX AC ACC72030;
XX
XX 08-JUL-2003 (first entry)
XX
XX BCU0571 gene #SEQ ID 37.
XX
```

```
XX Breast cancer; cytostatic; gene therapy; antisense therapy; regulated;
XX KW drug discovery; clinical medicine; forensic medicine; gene;
XX KW chromosome 8q24.12; ds.
XX
XX OS Homo sapiens.
XX
XX WO2003029421-A2.
XX
XX 10-APR-2003.
XX
XX 02-OCT-2002; 2002WO-US031287.
XX
XX 03-OCT-2001; 2001US-0326526P.
XX 14-MAY-2002; 2002US-00144194.
XX
XX (ORIG-) ORIGENE TECHNOLOGIES INC.
XX
XX Sun Z, Li X, Fan W, Kovacs KF, Jay G;
XX WPI; 2003-381623/36.
XX P-PSDB; ABR58296.
XX
XX New isolated human differentially-regulated breast cancer polynucleotide
XX PT and polypeptide, useful for diagnosing, staging, prognosticating,
XX PT preventing and/or treating diseases and conditions relating to breast
XX PT cancer.
XX
XX Claim 2; SEQ ID NO 37; 127pp + Sequence Listing; English.
XX
XX The invention relates to isolated polynucleotides which are
XX CC differentially-regulated in breast cancer. The methods and compositions
XX CC of the present invention are useful for detecting, diagnosing, staging,
XX CC monitoring, prognosticating, preventing and/or treating diseases and
XX CC conditions relating to breast cancer, and may be used in gene therapy or
XX CC antisense therapy. They can also be used in research, drug discovery,
XX CC clinical medicine and forensic medicine. Sequences given in records
XX CC ACC72012-ACC72074 represent polynucleotides of the invention that are
XX CC differentially-regulated in breast cancer. NOTE: The sequence data for
XX CC this patent did not form part of the printed specification, but was
XX CC obtained in electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences
XX
XX SQ Sequence 3288 BP; 967 A; 582 C; 753 G; 986 T; 0 U; 0 Other;
XX
XX Query Match 87.2%; Score 21.8; DB 8; Length 3288;
XX Best Local Similarity 92.0%; Pred. No. 11;
XX Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
XX QY 1 AAAAAAAAAATCAGGCGCAGGCATGG 25
XX ||||| ||||| ||||| ||||| |||||
XX Db 2979 AAAAAAAAAATCTTGGCCAGGCATGG 3003
XX
XX RESULT 8
XX AAK69927
XX ID AAK69927 standard; DNA; 5017 BP.
XX
XX AC AAK69927;
XX
XX 06-NOV-2001 (first entry)
XX
XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:24739.
XX
XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
XX KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX
XX OS Homo sapiens.
XX
XX WO200157182-A2.
XX
XX 09-AUG-2001.
XX
```


PF 17-JAN-2001; 2001WO-US001354.
 XX
 PR 31-JAN-2000; 2000US-0179065P.
 PR 04-FEB-2000; 2000US-0180828P.
 PR 24-FEB-2000; 2000US-0184664P.
 PR 02-MAR-2000; 2000US-0186350P.
 PR 16-MAR-2000; 2000US-0189874P.
 PR 17-MAR-2000; 2000US-0190076P.
 PR 18-APR-2000; 2000US-0198123P.
 PR 13-MAY-2000; 2000US-0205515P.
 PR 07-JUN-2000; 2000US-0209467P.
 PR 28-JUN-2000; 2000US-0214886P.
 PR 30-JUN-2000; 2000US-0215135P.
 PR 07-JUL-2000; 2000US-0216647P.
 PR 07-JUL-2000; 2000US-0216880P.
 PR 11-JUL-2000; 2000US-0217487P.
 PR 11-JUL-2000; 2000US-0217496P.
 PR 14-JUL-2000; 2000US-0218290P.
 PR 26-JUL-2000; 2000US-0220963P.
 PR 26-JUL-2000; 2000US-0220964P.
 PR 14-AUG-2000; 2000US-0224518P.
 PR 14-AUG-2000; 2000US-0224519P.
 PR 14-AUG-2000; 2000US-0225213P.
 PR 14-AUG-2000; 2000US-0225214P.
 PR 14-AUG-2000; 2000US-0225266P.
 PR 14-AUG-2000; 2000US-0225267P.
 PR 14-AUG-2000; 2000US-0225268P.
 PR 14-AUG-2000; 2000US-0225270P.
 PR 14-AUG-2000; 2000US-0225447P.
 PR 14-AUG-2000; 2000US-0225757P.
 PR 14-AUG-2000; 2000US-0225758P.
 PR 14-AUG-2000; 2000US-0225759P.
 PR 18-AUG-2000; 2000US-0226279P.
 PR 22-AUG-2000; 2000US-0226681P.
 PR 22-AUG-2000; 2000US-0226868P.
 PR 23-AUG-2000; 2000US-0227182P.
 PR 23-AUG-2000; 2000US-0227009P.
 PR 30-AUG-2000; 2000US-0228924P.
 PR 01-SEP-2000; 2000US-0229287P.
 PR 01-SEP-2000; 2000US-0229343P.
 PR 01-SEP-2000; 2000US-0229344P.
 PR 01-SEP-2000; 2000US-0229345P.
 PR 05-SEP-2000; 2000US-0229509P.
 PR 05-SEP-2000; 2000US-0229513P.
 PR 06-SEP-2000; 2000US-0230437P.
 PR 06-SEP-2000; 2000US-0230438P.
 PR 08-SEP-2000; 2000US-0231242P.
 PR 08-SEP-2000; 2000US-0231243P.
 PR 08-SEP-2000; 2000US-0231244P.
 PR 08-SEP-2000; 2000US-0231413P.
 PR 08-SEP-2000; 2000US-0231414P.
 PR 08-SEP-2000; 2000US-0232080P.
 PR 08-SEP-2000; 2000US-0232081P.
 PR 12-SEP-2000; 2000US-0231968P.
 PR 14-SEP-2000; 2000US-0232397P.
 PR 14-SEP-2000; 2000US-0232398P.
 PR 14-SEP-2000; 2000US-0232399P.
 PR 14-SEP-2000; 2000US-0232400P.
 PR 14-SEP-2000; 2000US-0232401P.
 PR 14-SEP-2000; 2000US-0233063P.
 PR 14-SEP-2000; 2000US-0233064P.
 PR 14-SEP-2000; 2000US-0233065P.
 PR 21-SEP-2000; 2000US-0234223P.
 PR 21-SEP-2000; 2000US-0234274P.
 PR 25-SEP-2000; 2000US-0234997P.
 PR 25-SEP-2000; 2000US-0234998P.
 PR 26-SEP-2000; 2000US-0235484P.
 PR 27-SEP-2000; 2000US-0235834P.
 PR 27-SEP-2000; 2000US-0235836P.
 PR 29-SEP-2000; 2000US-0236327P.
 PR 29-SEP-2000; 2000US-0236367P.
 PR 29-SEP-2000; 2000US-0236368P.
 PR 29-SEP-2000; 2000US-0236369P.

PR 29-SEP-2000; 2000US-0236370P.
 PR 02-OCT-2000; 2000US-0236802P.
 PR 02-OCT-2000; 2000US-0237037P.
 PR 02-OCT-2000; 2000US-0237038P.
 PR 02-OCT-2000; 2000US-0237039P.
 PR 02-OCT-2000; 2000US-0237040P.
 PR 13-OCT-2000; 2000US-0239935P.
 PR 13-OCT-2000; 2000US-0239937P.
 PR 20-OCT-2000; 2000US-0240960P.
 PR 20-OCT-2000; 2000US-0241221P.
 PR 20-OCT-2000; 2000US-0241785P.
 PR 20-OCT-2000; 2000US-0241786P.
 PR 20-OCT-2000; 2000US-0241787P.
 PR 20-OCT-2000; 2000US-0241808P.
 PR 20-OCT-2000; 2000US-0241809P.
 PR 20-OCT-2000; 2000US-0241826P.
 PR 01-NOV-2000; 2000US-0244617P.
 PR 08-NOV-2000; 2000US-0246474P.
 PR 08-NOV-2000; 2000US-0246475P.
 PR 08-NOV-2000; 2000US-0246476P.
 PR 08-NOV-2000; 2000US-0246477P.
 PR 08-NOV-2000; 2000US-0246478P.
 PR 08-NOV-2000; 2000US-0246523P.
 PR 08-NOV-2000; 2000US-0246524P.
 PR 08-NOV-2000; 2000US-0246525P.
 PR 08-NOV-2000; 2000US-0246526P.
 PR 08-NOV-2000; 2000US-0246527P.
 PR 08-NOV-2000; 2000US-0246528P.
 PR 08-NOV-2000; 2000US-0246532P.
 PR 08-NOV-2000; 2000US-0246609P.
 PR 08-NOV-2000; 2000US-0246610P.
 PR 08-NOV-2000; 2000US-0246611P.
 PR 08-NOV-2000; 2000US-0246613P.
 PR 17-NOV-2000; 2000US-0249207P.
 PR 17-NOV-2000; 2000US-0249208P.
 PR 17-NOV-2000; 2000US-0249209P.
 PR 17-NOV-2000; 2000US-0249210P.
 PR 17-NOV-2000; 2000US-0249211P.
 PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249244P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249264P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 17-NOV-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 05-DEC-2000; 2000US-0256719P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 08-DEC-2000; 2000US-0251990P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-02559678P.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-483426/52.

Nucleic acids encoding human immune/hematopoietic antigen polypeptides.

PT

PT useful for preventing, diagnosing and/or treating cancers and metastasis.
XX Disclosure; SEQ ID NO 24739; 3071pp + Sequence Listing; English.
XX

CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting the
CC nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
CC represent sequences used in the exemplification of the present invention
XX

SQ Sequence 5017 BP; 1338 A; 963 C; 1013 G; 1703 T; 0 U; 0 Other;

Query Match 87.2%; Score 21.8; DB 4; Length 5017;
Best Local Similarity 92.0%; Pred. No. 12;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAAAAAATCAGGCGCAGGCGATGG 25

||||||| ||||||| ||||||| |||||||
Db 3079 AAAAAAAAAAACGGCCAGGCGATGG 3103

RESULT 9

AAK69928

ID AAK69928 standard; DNA; 5017 BP.

XX

AC AAK69928;

XX

DT 06-NOV-2001 (first entry)

XX

DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:24740.

XX

XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;

KW cytostatic; gene therapy; vaccine; metastasis; ds.

XX

OS Homo sapiens.

XX

XX WO200157182-A2.

XX

PD 09-AUG-2001.

XX

XX 17-JAN-2001; 2001WO-US001354.

XX

XX 31-JAN-2000; 2000US-0179065P.

PR

PR 04-FEB-2000; 2000US-0180628P.

PR

PR 24-FEB-2000; 2000US-0184664P.

PR

PR 02-MAR-2000; 2000US-0186350P.

PR

PR 16-MAR-2000; 2000US-0189874P.

PR

PR 17-MAR-2000; 2000US-0190076P.

PR

PR 18-APR-2000; 2000US-0198123P.

PR

PR 19-MAY-2000; 2000US-0205515P.

PR

PR 28-JUN-2000; 2000US-0209467P.

PR

PR 30-JUN-2000; 2000US-0214886P.

PR

PR 07-JUL-2000; 2000US-0215135P.

PR

PR 07-JUL-2000; 2000US-0216647P.

PR

PR 07-JUL-2000; 2000US-0216880P.

PR

PR 11-JUL-2000; 2000US-0217487P.

PR

PR 11-JUL-2000; 2000US-0217496P.

PR

PR 14-JUL-2000; 2000US-0218290P.

PR

PR 26-JUL-2000; 2000US-0220963P.

PR

PR 26-JUL-2000; 2000US-0220964P.

PR

PR 14-AUG-2000; 2000US-0224518P.

PR

PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.

WP Sequence split into 4 fragments LOCUS ACN45090 Accession ACN45090
WP Fragment Name Begin End
WP ACN45090_0 1 110000
WP ACN45090_1 100001 210000
WP ACN45090_2 200001 310000
WP ACN45090_3 300001 350764
ID ACN45090 standard; DNA; 350764 BP.
XX
XX
AC ACN45090;
XX

DT 18-NOV-2004 (first entry)
XX
XX Human genomic sequence hCG22125.
DE
XX Cytostatic; carcinoma; lymphoma; cancer; human; gene; ss.
KW
XX Homo sapiens.
XX WO2003073826-A2.
XX
XX 12-SEP-2003.
XX

XX 28-FEB-2003; 2003WO-US06235.
XX
XX 01-MAR-2002; 2002US-00087192.
XX (SAGR-) SAGRES DISCOVERY.
PA
XX Morris DW;
XX
XX WPI; 2003-328604/31.
XX

PT Recombinant nucleic acid useful for diagnosis and treatment of carcinoma
PT comprises a nucleotide sequence.
XX
XX Claim 1; SEQ ID NO 1864; Opp; English.
XX

CC The present invention relates to novel DNA and protein sequences which
CC are associated with carcinomas. The sequences are useful for: (i) for
CC screening drug candidates; (ii) for screening of bioactive agent capable
CC of binding to Carcinoma Associated Protein (CAP); (iii) for screening of
CC a bioactive agent capable of modulating the activity of CAP; (iv) for
CC evaluating the effect of a candidate carcinoma drug; (v) for diagnosing
CC carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating
CC carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip;
CC (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for
CC determining Carcinoma Associated (CA) gene copy number. In addition, the
CC CA genes are useful as DNA vaccines and the CAP are useful as markers of
CC carcinoma including lymphoma. The present sequence is one such CA coding
CC sequence. Note: This patent is an equivalent to basic patent
CC US2002182586A1, for which no sequence data was published
XX
SQ Sequence 350764 BP; 99569 A; 63584 C; 68015 G; 119596 T; 0 U; 0 Other;

Query Match 87.2%; Score 21.8; DB 11; Length 110000;
Best Local Similarity 92.0%; Pred. No. 20;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 AAAAAAAAAATCAGGCCAGGCTGG 25
|||||||
Db 23134 AAAAAAAAAAAGCCAGGCTGG 23110
|||||||

RESULT 12
AAK79671
ID AAK79671 standard; DNA; 25574 BP.
XX
XX AAK79671;
AC
XX 07-NOV-2001 (first entry)
DT
XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:34483.
DE
XX

KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX
OS Homo sapiens.
XX WO200157182-A2.
XX
XX 09-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US001354.
XX
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.

25-SEP-2000; 2000US-0234997P.
 PR 25-SEP-2000; 2000US-0234998P.
 PR 26-SEP-2000; 2000US-0235484P.
 PR 27-SEP-2000; 2000US-0235834P.
 PR 27-SEP-2000; 2000US-0235836P.
 PR 28-SEP-2000; 2000US-0236327P.
 PR 29-SEP-2000; 2000US-0236367P.
 PR 29-SEP-2000; 2000US-0236368P.
 PR 29-SEP-2000; 2000US-0236369P.
 PR 29-SEP-2000; 2000US-0236370P.
 PR 02-OCT-2000; 2000US-0236802P.
 PR 02-OCT-2000; 2000US-0237037P.
 PR 02-OCT-2000; 2000US-0237038P.
 PR 02-OCT-2000; 2000US-0237039P.
 PR 02-OCT-2000; 2000US-0237040P.
 PR 13-OCT-2000; 2000US-0239935P.
 PR 13-OCT-2000; 2000US-0239937P.
 PR 20-OCT-2000; 2000US-0240960P.
 PR 20-OCT-2000; 2000US-0241221P.
 PR 20-OCT-2000; 2000US-0241785P.
 PR 20-OCT-2000; 2000US-0241786P.
 PR 20-OCT-2000; 2000US-0241787P.
 PR 20-OCT-2000; 2000US-0241808P.
 PR 20-OCT-2000; 2000US-0241809P.
 PR 01-NOV-2000; 2000US-0241826P.
 PR 08-NOV-2000; 2000US-0246474P.
 PR 08-NOV-2000; 2000US-0246475P.
 PR 08-NOV-2000; 2000US-0246476P.
 PR 08-NOV-2000; 2000US-0246477P.
 PR 08-NOV-2000; 2000US-0246478P.
 PR 08-NOV-2000; 2000US-0246523P.
 PR 08-NOV-2000; 2000US-0246524P.
 PR 08-NOV-2000; 2000US-0246525P.
 PR 08-NOV-2000; 2000US-0246526P.
 PR 08-NOV-2000; 2000US-0246527P.
 PR 08-NOV-2000; 2000US-0246528P.
 PR 08-NOV-2000; 2000US-0246532P.
 PR 08-NOV-2000; 2000US-0246609P.
 PR 08-NOV-2000; 2000US-0246610P.
 PR 08-NOV-2000; 2000US-0246611P.
 PR 08-NOV-2000; 2000US-0246613P.
 PR 17-NOV-2000; 2000US-0249207P.
 PR 17-NOV-2000; 2000US-0249208P.
 PR 17-NOV-2000; 2000US-0249209P.
 PR 17-NOV-2000; 2000US-0249210P.
 PR 17-NOV-2000; 2000US-0249211P.
 PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249244P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249264P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 05-DEC-2000; 2000US-0256719P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 08-DEC-2000; 2000US-0251990P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA Rosen CA, Barash SC, Ruben SM;
 PI WPI; 2001-483426/52.
 XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating cancers and metastasis.
 XX Disclosure; SEQ ID NO 34483; 307lpp + Sequence Listing; English.
 PS AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
 CC amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic
 CC activity, and can be used in gene therapy and vaccine production. (I)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patients own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I), by inserting the
 CC nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/haematopoietic-related diseases, especially
 CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
 CC to AAK87694 represent human immune/haematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
 CC represent sequences used in the exemplification of the present invention
 XX SQ Sequence 25574 BP; 7154 A; 6018 C; 6058 G; 6344 T; 0 U; 0 Other;
 Query Match 83.2%; Score 20.8; DB 4; Length 25574;
 Best Local Similarity 91.7%; Pred. No. 44;
 Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 AAAAAAAAAATCAGCGCCAGCGCATG 24
 |||||
 DB 12227 AAAAAAAAAATCAGCGCCAGCGCATG 12250
 RESULT 13
 AAK83760
 ID AAK83760 standard; DNA; 25574 BP.
 XX AAK83760;
 AC AAK83760;
 XX 07-NOV-2001 (first entry)
 DT Human immune/haematopoietic antigen genomic sequence SEQ ID NO:38572.
 XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
 KW cytostatic; gene therapy; vaccine; metastasis; ds.
 XX Homo sapiens.
 OS WO200157182-A2.
 XX 09-AUG-2001.
 PD 17-JAN-2001; 2001WO-US001354.
 PF 31-JAN-2000; 2000US-0179065P.
 PR 04-FEB-2000; 2000US-0180628P.
 PR 24-FEB-2000; 2000US-0184664P.
 PR 02-MAR-2000; 2000US-0186350P.
 PR 16-MAR-2000; 2000US-0189874P.
 PR 17-MAR-2000; 2000US-0190076P.
 PR 18-APR-2000; 2000US-0198123P.
 PR 19-MAY-2000; 2000US-0205515P.
 PR 07-JUN-2000; 2000US-0209467P.
 PR 28-JUN-2000; 2000US-0214886P.

PR	30-JUN-2000;	2000US-0215135P.	PR	20-OCT-2000;	2000US-0241787P.
PR	07-JUL-2000;	2000US-0216647P.	PR	20-OCT-2000;	2000US-0241808P.
PR	07-JUL-2000;	2000US-0216880P.	PR	20-OCT-2000;	2000US-0241809P.
PR	11-JUL-2000;	2000US-0217480P.	PR	20-OCT-2000;	2000US-0241826P.
PR	11-JUL-2000;	2000US-0217496P.	PR	01-NOV-2000;	2000US-0244617P.
PR	14-JUL-2000;	2000US-0218290P.	PR	08-NOV-2000;	2000US-0246474P.
PR	26-JUL-2000;	2000US-0220963P.	PR	08-NOV-2000;	2000US-0246475P.
PR	26-JUL-2000;	2000US-0220964P.	PR	08-NOV-2000;	2000US-0246476P.
PR	14-AUG-2000;	2000US-0224518P.	PR	08-NOV-2000;	2000US-0246477P.
PR	14-AUG-2000;	2000US-0224519P.	PR	08-NOV-2000;	2000US-0246478P.
PR	14-AUG-2000;	2000US-0225213P.	PR	08-NOV-2000;	2000US-0246523P.
PR	14-AUG-2000;	2000US-0225214P.	PR	08-NOV-2000;	2000US-0246524P.
PR	14-AUG-2000;	2000US-0225266P.	PR	08-NOV-2000;	2000US-0246525P.
PR	14-AUG-2000;	2000US-0225267P.	PR	08-NOV-2000;	2000US-0246526P.
PR	14-AUG-2000;	2000US-0225268P.	PR	08-NOV-2000;	2000US-0246527P.
PR	14-AUG-2000;	2000US-0225269P.	PR	08-NOV-2000;	2000US-0246528P.
PR	14-AUG-2000;	2000US-0225270P.	PR	08-NOV-2000;	2000US-0246532P.
PR	14-AUG-2000;	2000US-0225447P.	PR	08-NOV-2000;	2000US-0246609P.
PR	14-AUG-2000;	2000US-0225757P.	PR	08-NOV-2000;	2000US-0246610P.
PR	14-AUG-2000;	2000US-0225758P.	PR	08-NOV-2000;	2000US-0246611P.
PR	14-AUG-2000;	2000US-0225759P.	PR	08-NOV-2000;	2000US-0246613P.
PR	18-AUG-2000;	2000US-0226279P.	PR	17-NOV-2000;	2000US-0249207P.
PR	22-AUG-2000;	2000US-0226681P.	PR	17-NOV-2000;	2000US-0249208P.
PR	22-AUG-2000;	2000US-0226868P.	PR	17-NOV-2000;	2000US-0249209P.
PR	22-AUG-2000;	2000US-0227182P.	PR	17-NOV-2000;	2000US-0249210P.
PR	23-AUG-2000;	2000US-0227009P.	PR	17-NOV-2000;	2000US-0249211P.
PR	30-AUG-2000;	2000US-0228924P.	PR	17-NOV-2000;	2000US-0249212P.
PR	01-SEP-2000;	2000US-0229287P.	PR	17-NOV-2000;	2000US-0249213P.
PR	01-SEP-2000;	2000US-0229343P.	PR	17-NOV-2000;	2000US-0249214P.
PR	01-SEP-2000;	2000US-0229344P.	PR	17-NOV-2000;	2000US-0249215P.
PR	01-SEP-2000;	2000US-0229345P.	PR	17-NOV-2000;	2000US-0249216P.
PR	05-SEP-2000;	2000US-0229509P.	PR	17-NOV-2000;	2000US-0249217P.
PR	05-SEP-2000;	2000US-0229513P.	PR	17-NOV-2000;	2000US-0249218P.
PR	06-SEP-2000;	2000US-0230437P.	PR	17-NOV-2000;	2000US-0249244P.
PR	06-SEP-2000;	2000US-0230438P.	PR	17-NOV-2000;	2000US-0249245P.
PR	08-SEP-2000;	2000US-0231242P.	PR	17-NOV-2000;	2000US-0249264P.
PR	08-SEP-2000;	2000US-0231243P.	PR	17-NOV-2000;	2000US-0249265P.
PR	08-SEP-2000;	2000US-0231244P.	PR	17-NOV-2000;	2000US-0249297P.
PR	08-SEP-2000;	2000US-0231413P.	PR	17-NOV-2000;	2000US-0249299P.
PR	08-SEP-2000;	2000US-0231414P.	PR	17-NOV-2000;	2000US-0249300P.
PR	08-SEP-2000;	2000US-0232080P.	PR	01-DEC-2000;	2000US-0250160P.
PR	08-SEP-2000;	2000US-0232081P.	PR	01-DEC-2000;	2000US-0250391P.
PR	12-SEP-2000;	2000US-0231968P.	PR	05-DEC-2000;	2000US-0251030P.
PR	14-SEP-2000;	2000US-0232397P.	PR	05-DEC-2000;	2000US-0251988P.
PR	14-SEP-2000;	2000US-0232398P.	PR	05-DEC-2000;	2000US-0256719P.
PR	14-SEP-2000;	2000US-0232399P.	PR	06-DEC-2000;	2000US-0251479P.
PR	14-SEP-2000;	2000US-0232400P.	PR	08-DEC-2000;	2000US-0251856P.
PR	14-SEP-2000;	2000US-0232401P.	PR	08-DEC-2000;	2000US-0251868P.
PR	14-SEP-2000;	2000US-0233063P.	PR	08-DEC-2000;	2000US-0251869P.
PR	14-SEP-2000;	2000US-0233064P.	PR	08-DEC-2000;	2000US-0251989P.
PR	21-SEP-2000;	2000US-0234223P.	PR	08-DEC-2000;	2000US-0251990P.
PR	21-SEP-2000;	2000US-0234274P.	PR	11-DEC-2000;	2000US-0254097P.
PR	25-SEP-2000;	2000US-0234997P.	PR	05-JAN-2001;	2001US-0259678P.
PR	25-SEP-2000;	2000US-0234998P.	XX		
PR	26-SEP-2000;	2000US-0235484P.	XX		
PR	27-SEP-2000;	2000US-0235834P.	XX		
PR	27-SEP-2000;	2000US-0235836P.	XX		
PR	29-SEP-2000;	2000US-0236327P.	XX		
PR	29-SEP-2000;	2000US-0236367P.	XX		
PR	29-SEP-2000;	2000US-0236368P.	XX		
PR	29-SEP-2000;	2000US-0236369P.	XX		
PR	29-SEP-2000;	2000US-0236370P.	XX		
PR	02-OCT-2000;	2000US-0237037P.	XX		
PR	02-OCT-2000;	2000US-0237038P.	XX		
PR	02-OCT-2000;	2000US-0237039P.	XX		
PR	02-OCT-2000;	2000US-0237040P.	XX		
PR	13-OCT-2000;	2000US-0239935P.	XX		
PR	13-OCT-2000;	2000US-0239937P.	XX		
PR	20-OCT-2000;	2000US-0240960P.	XX		
PR	20-OCT-2000;	2000US-0241221P.	XX		
PR	20-OCT-2000;	2000US-0241785P.	XX		
PR	20-OCT-2000;	2000US-0241786P.	XX		

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-483426/52.

Nucleic acids encoding human immune/hematopoietic antigen polypeptides, useful for preventing, diagnosing and/or treating cancers and metastasis.

Disclosure; SEQ ID NO 38572; 307lpp + Sequence Listing; English.

AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I) amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic activity, and can be used in gene therapy and vaccine production. (I) proteins, and polynucleotides may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate (I) expression. For example, they may be used to treat disorders associated with decreased expression by rectifying mutations or deletions in a patient's genome that affect the activity of (I) by expressing inactive proteins or to

CC supplement the patients own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I), by inserting the
 CC nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/haematopoietic-related diseases, especially
 CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
 CC to AAK87694 represent human immune/haematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
 CC represent sequences used in the exemplification of the present invention
 XX
 SQ Sequence 25574 BP; 7154 A; 6018 C; 6058 G; 6344 T; 0 U; 0 Other;

Query Match 83.2%; Score 20.8; DB 4; Length 25574;
 Best Local Similarity 91.7%; Pred. No. 44;
 Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AAAAAAAAAATCAGCGCCAGGCATG 24
 |||||
 Db 12227 AAAAAAAAAATCAGCGCCAGGCATG 12250

RESULT 14
 AAL05619
 ID AAL05619 standard; DNA; 25574 BP.

XX AC AAL05619;

XX DT 21-NOV-2001 (first entry)

XX DE Human reproductive system related antigen DNA SEQ ID NO: 8307.

XX KW Human; reproductive system related antigen; reproductive system disorder;
 XX cancer; gene therapy; ds.

XX OS Homo sapiens.

XX PN WO200155320-A2.

XX PD 02-AUG-2001.

XX PF 17-JAN-2001; 2001WO-US001339.

XX PR 31-JAN-2000; 2000US-0179065P.

XX PR 04-FEB-2000; 2000US-0180628P.

XX PR 24-FEB-2000; 2000US-0184664P.

XX PR 02-MAR-2000; 2000US-0186350P.

XX PR 16-MAR-2000; 2000US-0189874P.

XX PR 17-MAR-2000; 2000US-0190076P.

XX PR 18-APR-2000; 2000US-0198123P.

XX PR 19-MAY-2000; 2000US-0205515P.

XX PR 07-JUN-2000; 2000US-0209467P.

XX PR 28-JUN-2000; 2000US-0214886P.

XX PR 30-JUN-2000; 2000US-0215135P.

XX PR 07-JUL-2000; 2000US-0216647P.

XX PR 07-JUL-2000; 2000US-0216880P.

XX PR 11-JUL-2000; 2000US-0217487P.

XX PR 11-JUL-2000; 2000US-0217496P.

XX PR 14-JUL-2000; 2000US-0218290P.

XX PR 26-JUL-2000; 2000US-0220963P.

XX PR 26-JUL-2000; 2000US-0220964P.

XX PR 14-AUG-2000; 2000US-0224518P.

XX PR 14-AUG-2000; 2000US-0224519P.
 XX PR 14-AUG-2000; 2000US-0225213P.
 XX PR 14-AUG-2000; 2000US-0225214P.
 XX PR 14-AUG-2000; 2000US-0225266P.
 XX PR 14-AUG-2000; 2000US-0225267P.
 XX PR 14-AUG-2000; 2000US-0225268P.
 XX PR 14-AUG-2000; 2000US-0225270P.
 XX PR 14-AUG-2000; 2000US-0225447P.
 XX PR 14-AUG-2000; 2000US-0225757P.
 XX PR 14-AUG-2000; 2000US-0225758P.
 XX PR 14-AUG-2000; 2000US-0225759P.
 XX PR 18-AUG-2000; 2000US-0226279P.
 XX PR 31-JAN-2000; 2000US-0179065P.
 XX PR 04-FEB-2000; 2000US-0180628P.
 XX PR 24-FEB-2000; 2000US-0184664P.
 XX PR 02-MAR-2000; 2000US-0186350P.
 XX PR 16-MAR-2000; 2000US-0189874P.
 XX PR 17-MAR-2000; 2000US-0190076P.
 XX PR 18-APR-2000; 2000US-0198123P.
 XX PR 19-MAY-2000; 2000US-0205515P.
 XX PR 07-JUN-2000; 2000US-0209467P.
 XX PR 28-JUN-2000; 2000US-0214886P.
 XX PR 30-JUN-2000; 2000US-0215135P.
 XX PR 07-JUL-2000; 2000US-0216647P.
 XX PR 07-JUL-2000; 2000US-0216880P.
 XX PR 11-JUL-2000; 2000US-0217487P.
 XX PR 11-JUL-2000; 2000US-0217496P.
 XX PR 14-JUL-2000; 2000US-0218290P.
 XX PR 26-JUL-2000; 2000US-0220963P.
 XX PR 26-JUL-2000; 2000US-0220964P.
 XX PR 14-AUG-2000; 2000US-0224518P.
 XX PR 14-AUG-2000; 2000US-0224519P.
 XX PR 14-AUG-2000; 2000US-0225213P.
 XX PR 14-AUG-2000; 2000US-0225214P.
 XX PR 14-AUG-2000; 2000US-0225266P.
 XX PR 14-AUG-2000; 2000US-0225267P.
 XX PR 14-AUG-2000; 2000US-0225268P.
 XX PR 14-AUG-2000; 2000US-0225270P.
 XX PR 14-AUG-2000; 2000US-0225447P.
 XX PR 14-AUG-2000; 2000US-0225757P.
 XX PR 14-AUG-2000; 2000US-0225758P.
 XX PR 14-AUG-2000; 2000US-0225759P.
 XX PR 18-AUG-2000; 2000US-0226279P.
 XX PR 22-AUG-2000; 2000US-0226681P.
 XX PR 22-AUG-2000; 2000US-0226688P.
 XX PR 23-AUG-2000; 2000US-0227182P.
 XX PR 30-AUG-2000; 2000US-0227009P.
 XX PR 01-SEP-2000; 2000US-0228924P.
 XX PR 01-SEP-2000; 2000US-0229287P.
 XX PR 01-SEP-2000; 2000US-0229343P.
 XX PR 01-SEP-2000; 2000US-0229344P.
 XX PR 05-SEP-2000; 2000US-0229345P.
 XX PR 05-SEP-2000; 2000US-0229509P.
 XX PR 05-SEP-2000; 2000US-0229513P.
 XX PR 06-SEP-2000; 2000US-0230437P.
 XX PR 06-SEP-2000; 2000US-0230438P.
 XX PR 08-SEP-2000; 2000US-0231242P.
 XX PR 08-SEP-2000; 2000US-0231243P.
 XX PR 08-SEP-2000; 2000US-0231244P.
 XX PR 08-SEP-2000; 2000US-0231413P.
 XX PR 08-SEP-2000; 2000US-0231414P.
 XX PR 08-SEP-2000; 2000US-0232080P.
 XX PR 08-SEP-2000; 2000US-0232081P.
 XX PR 12-SEP-2000; 2000US-0231968P.
 XX PR 14-SEP-2000; 2000US-0232397P.
 XX PR 14-SEP-2000; 2000US-0232398P.
 XX PR 14-SEP-2000; 2000US-0232399P.
 XX PR 14-SEP-2000; 2000US-0232400P.
 XX PR 14-SEP-2000; 2000US-0232401P.
 XX PR 14-SEP-2000; 2000US-0233063P.
 XX PR 14-SEP-2000; 2000US-0233064P.
 XX PR 14-SEP-2000; 2000US-0233065P.
 XX PR 21-SEP-2000; 2000US-0234223P.
 XX PR 21-SEP-2000; 2000US-0234274P.
 XX PR 25-SEP-2000; 2000US-0234997P.
 XX PR 25-SEP-2000; 2000US-0234998P.
 XX PR 26-SEP-2000; 2000US-0235484P.
 XX PR 27-SEP-2000; 2000US-0235834P.
 XX PR 27-SEP-2000; 2000US-0235836P.
 XX PR 29-SEP-2000; 2000US-0236327P.
 XX PR 29-SEP-2000; 2000US-0236367P.
 XX PR 29-SEP-2000; 2000US-0236368P.
 XX PR 29-SEP-2000; 2000US-0236369P.
 XX PR 29-SEP-2000; 2000US-0236370P.
 XX PR 29-SEP-2000; 2000US-0236802P.
 XX PR 02-OCT-2000; 2000US-0237037P.
 XX PR 02-OCT-2000; 2000US-0237038P.
 XX PR 02-OCT-2000; 2000US-0237039P.
 XX PR 02-OCT-2000; 2000US-0237040P.
 XX PR 13-OCT-2000; 2000US-0239935P.
 XX PR 13-OCT-2000; 2000US-0239937P.
 XX PR 20-OCT-2000; 2000US-0240960P.
 XX PR 20-OCT-2000; 2000US-0241321P.
 XX PR 20-OCT-2000; 2000US-0241785P.
 XX PR 20-OCT-2000; 2000US-0241786P.
 XX PR 20-OCT-2000; 2000US-0241787P.
 XX PR 20-OCT-2000; 2000US-0241808P.
 XX PR 20-OCT-2000; 2000US-0241809P.
 XX PR 20-OCT-2000; 2000US-0241826P.
 XX PR 01-NOV-2000; 2000US-0244617P.
 XX PR 08-NOV-2000; 2000US-0246474P.
 XX PR 08-NOV-2000; 2000US-0246475P.
 XX PR 08-NOV-2000; 2000US-0246476P.
 XX PR 08-NOV-2000; 2000US-0246477P.
 XX PR 08-NOV-2000; 2000US-0246478P.
 XX PR 08-NOV-2000; 2000US-0246523P.
 XX PR 08-NOV-2000; 2000US-0246524P.
 XX PR 08-NOV-2000; 2000US-0246525P.
 XX PR 08-NOV-2000; 2000US-0246526P.
 XX PR 08-NOV-2000; 2000US-0246527P.
 XX PR 08-NOV-2000; 2000US-0246528P.
 XX PR 08-NOV-2000; 2000US-0246532P.
 XX PR 08-NOV-2000; 2000US-0246609P.
 XX PR 08-NOV-2000; 2000US-0246610P.
 XX PR 08-NOV-2000; 2000US-0246611P.
 XX PR 08-NOV-2000; 2000US-0246613P.

PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
PI
XX
XX WPI; 2001-465570/50.
XX
PT Isolated nucleic acid molecule encoding a reproductive system antigen is
used in preventing, treating or ameliorating a medical condition.
XX
XX Disclosure; SEQ ID NO 8307; 1297pp + Sequence Listing; English.
XX
CC The present invention provides the protein and coding sequences of a
number of human reproductive system related antigens. These can be used
in the prevention and treatment of reproductive system disorders,
CC including cancer. The present sequence is a genomic sequence encoding a
protein of the invention
XX
XX Sequence 25574 BP; 7154 A; 6018 C; 6058 G; 6344 T; 0 U; 0 Other;
SQ
Query Match 83.2%; Score 20.8; DB 4; Length 25574;
Best Local Similarity 91.7%; Pred. No. 44;
Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 AAAAAAAAAATCAGCGCAGGCGATG 24
|||||||
Db 12227 AAAAAAAAAATCAGCGCAGGCGTG 12250
RESULT 15
AAK85305
ID AAK85305 standard; DNA; 25576 BP.
XX
AC AAK85305;
XX
DT 07-NOV-2001 (first entry)
XX
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:40117.
XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
KW

XX
OS Homo sapiens.
XX WO200157182-A2.
XX
XX 09-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US001354.
XX
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0228287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 25-SEP-2000; 2000US-0234999P.


```
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 06-DEC-2000; 2000US-0256719P.
PR 08-DEC-2000; 2000US-0251478P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 11-DEC-2000; 2000US-0251990P.
PR 05-JAN-2001; 2001US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
(HUMA-) HUMAN GENOME SCI INC.
PA
XX Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-483426/52.
XX
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and metastasis.
XX
XX Disclosure; SEQ ID NO 40117; 3071pp + Sequence Listing; English.
XX
XX AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patient's own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting the
CC nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC represent sequences used in the exemplification of the present invention
XX
XX Sequence 25576 BP; 7153 A; 6020 C; 6060 G; 6343 T; 0 U; 0 Other;
SQ
Query Match 83.2%; Score 20.8; DB 4; Length 25576;
Best Local Similarity 91.7%; Pred. No. 44;
Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 AAAAAAAAAATCAGCGCCAGCGCATG 24
|||||
Db 12225 AAAAAAAAAATCAAGCGCCAGCGGTG 12248
RESULT 16
AAK79669
ID AAK79669 standard; DNA; 25576 BP.
XX
AC AAK79669;
XX
XX 07-NOV-2001 (first entry)
DT
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:34481.
XX
XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX
XX Homo sapiens.
XX
XX WO200157182-A2.
XX
XX 09-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US001354.
XX
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
XX
```

PR 07-JUL-2000; 2000US-0216880P.
 PR 11-JUL-2000; 2000US-0217487P.
 PR 11-JUL-2000; 2000US-0217496P.
 PR 14-JUL-2000; 2000US-0218290P.
 PR 26-JUL-2000; 2000US-0220963P.
 PR 26-JUL-2000; 2000US-0220964P.
 PR 14-AUG-2000; 2000US-0224518P.
 PR 14-AUG-2000; 2000US-0224519P.
 PR 14-AUG-2000; 2000US-0225213P.
 PR 14-AUG-2000; 2000US-0225214P.
 PR 14-AUG-2000; 2000US-0225266P.
 PR 14-AUG-2000; 2000US-0225267P.
 PR 14-AUG-2000; 2000US-0225268P.
 PR 14-AUG-2000; 2000US-0225270P.
 PR 14-AUG-2000; 2000US-0225447P.
 PR 14-AUG-2000; 2000US-0225477P.
 PR 14-AUG-2000; 2000US-0225757P.
 PR 14-AUG-2000; 2000US-0225758P.
 PR 14-AUG-2000; 2000US-0225759P.
 PR 18-AUG-2000; 2000US-0226279P.
 PR 22-AUG-2000; 2000US-0226581P.
 PR 22-AUG-2000; 2000US-0226868P.
 PR 22-AUG-2000; 2000US-0227182P.
 PR 23-AUG-2000; 2000US-0227009P.
 PR 30-AUG-2000; 2000US-0228924P.
 PR 01-SEP-2000; 2000US-0229287P.
 PR 01-SEP-2000; 2000US-0229343P.
 PR 01-SEP-2000; 2000US-0229344P.
 PR 01-SEP-2000; 2000US-0229345P.
 PR 05-SEP-2000; 2000US-0229509P.
 PR 05-SEP-2000; 2000US-0229513P.
 PR 06-SEP-2000; 2000US-0230437P.
 PR 06-SEP-2000; 2000US-0230438P.
 PR 08-SEP-2000; 2000US-0231142P.
 PR 08-SEP-2000; 2000US-0231143P.
 PR 08-SEP-2000; 2000US-0231144P.
 PR 08-SEP-2000; 2000US-0231145P.
 PR 08-SEP-2000; 2000US-0232080P.
 PR 08-SEP-2000; 2000US-0232081P.
 PR 12-SEP-2000; 2000US-0231968P.
 PR 14-SEP-2000; 2000US-0232397P.
 PR 14-SEP-2000; 2000US-0232398P.
 PR 14-SEP-2000; 2000US-0232399P.
 PR 14-SEP-2000; 2000US-0232400P.
 PR 14-SEP-2000; 2000US-0232401P.
 PR 14-SEP-2000; 2000US-0233063P.
 PR 14-SEP-2000; 2000US-0233064P.
 PR 14-SEP-2000; 2000US-0233065P.
 PR 21-SEP-2000; 2000US-0234223P.
 PR 21-SEP-2000; 2000US-0234274P.
 PR 25-SEP-2000; 2000US-0234997P.
 PR 25-SEP-2000; 2000US-0234998P.
 PR 26-SEP-2000; 2000US-0235484P.
 PR 27-SEP-2000; 2000US-0235834P.
 PR 27-SEP-2000; 2000US-0235836P.
 PR 29-SEP-2000; 2000US-0236327P.
 PR 29-SEP-2000; 2000US-0236367P.
 PR 29-SEP-2000; 2000US-0236368P.
 PR 29-SEP-2000; 2000US-0236369P.
 PR 29-SEP-2000; 2000US-0236370P.
 PR 02-OCT-2000; 2000US-0236802P.
 PR 02-OCT-2000; 2000US-0237037P.
 PR 02-OCT-2000; 2000US-0237038P.
 PR 02-OCT-2000; 2000US-0237039P.
 PR 13-OCT-2000; 2000US-0237040P.
 PR 13-OCT-2000; 2000US-0239335P.
 PR 13-OCT-2000; 2000US-0239337P.
 PR 20-OCT-2000; 2000US-0240960P.
 PR 20-OCT-2000; 2000US-0241221P.
 PR 20-OCT-2000; 2000US-0241785P.
 PR 20-OCT-2000; 2000US-0241786P.
 PR 20-OCT-2000; 2000US-0241787P.
 PR 20-OCT-2000; 2000US-0241808P.

PR 20-OCT-2000; 2000US-0241809P.
 PR 20-OCT-2000; 2000US-0241826P.
 PR 01-NOV-2000; 2000US-0244617P.
 PR 08-NOV-2000; 2000US-0246474P.
 PR 08-NOV-2000; 2000US-0246475P.
 PR 08-NOV-2000; 2000US-0246476P.
 PR 08-NOV-2000; 2000US-0246477P.
 PR 08-NOV-2000; 2000US-0246478P.
 PR 08-NOV-2000; 2000US-0246523P.
 PR 08-NOV-2000; 2000US-0246524P.
 PR 08-NOV-2000; 2000US-0246525P.
 PR 08-NOV-2000; 2000US-0246526P.
 PR 08-NOV-2000; 2000US-0246527P.
 PR 08-NOV-2000; 2000US-0246528P.
 PR 08-NOV-2000; 2000US-0246532P.
 PR 08-NOV-2000; 2000US-0246609P.
 PR 08-NOV-2000; 2000US-0246610P.
 PR 08-NOV-2000; 2000US-0246611P.
 PR 08-NOV-2000; 2000US-0246613P.
 PR 17-NOV-2000; 2000US-0249207P.
 PR 17-NOV-2000; 2000US-0249208P.
 PR 17-NOV-2000; 2000US-0249209P.
 PR 17-NOV-2000; 2000US-0249210P.
 PR 17-NOV-2000; 2000US-0249211P.
 PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249244P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249264P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 05-DEC-2000; 2000US-0256719P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 08-DEC-2000; 2000US-0251990P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-02559678P.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-483426/52.

Nucleic acids encoding human immune/hematopoietic antigen polypeptides, useful for preventing, diagnosing and/or treating cancers and metastasis.

Disclosure; SEQ ID NO 34481; 3071pp + Sequence Listing; English.

AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I) amino acid sequences given in AAK62170 to AAK61921. (I) have cytostatic activity, and can be used in gene therapy and vaccine production. (I) proteins and polynucleotides may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate (I) expression. For example, they may be used to treat disorders associated with decreased expression by rectifying mutations or deletions in a patient's genome that affect the activity of (I) by expressing inactive proteins or to supplement the patients own production of (I). Additionally, (I) polynucleotides may be used to produce the secreted (I), by inserting the

CC nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC represent sequences used in the exemplification of the present invention

SQ Sequence 25576 BP; 7153 A; 6020 C; 6060 G; 6343 T; 0 U; 0 Other;

Query Match 83.2%; Score 20.8; DB 4; Length 25576;
Best Local Similarity 91.7%; Pred. No. 44;
Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AAAAAAAAAACAGCCGAGCATG 24

Db 12225 AAAAAAAAAACAGCCGAGCGTG 12248

RESULT 17

AAK83758
ID AAK83758 standard; DNA; 25576 BP.

AC AAK83758;

DT 07-NOV-2001 (first entry)

XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:38570.

DE Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.

XX Homo sapiens.

XX WO200157182-A2.

XX 09-AUG-2001.

XX 17-JAN-2001; 2001WO-US001354.

XX 31-JAN-2000; 2000US-0179065P.

PR 04-FEB-2000; 2000US-0180628P.

PR 24-FEB-2000; 2000US-0184664P.

PR 02-MAR-2000; 2000US-0186350P.

PR 16-MAR-2000; 2000US-0189874P.

PR 17-MAR-2000; 2000US-0190076P.

PR 18-APR-2000; 2000US-0198123P.

PR 19-MAY-2000; 2000US-0205515P.

PR 07-JUN-2000; 2000US-0209467P.

PR 28-JUN-2000; 2000US-0214886P.

PR 30-JUN-2000; 2000US-0215135P.

PR 07-JUL-2000; 2000US-0216647P.

PR 07-JUL-2000; 2000US-0216880P.

PR 11-JUL-2000; 2000US-0217487P.

PR 11-JUL-2000; 2000US-0217496P.

PR 14-JUL-2000; 2000US-0218290P.

PR 26-JUL-2000; 2000US-0220963P.

PR 26-JUL-2000; 2000US-0220964P.

PR 14-AUG-2000; 2000US-0224518P.

PR 14-AUG-2000; 2000US-0224519P.

PR 14-AUG-2000; 2000US-0225211P.

PR 14-AUG-2000; 2000US-0225214P.

PR 14-AUG-2000; 2000US-0225266P.

PR 14-AUG-2000; 2000US-0225267P.

PR 14-AUG-2000; 2000US-0225268P.

PR 14-AUG-2000; 2000US-0225270P.

PR 14-AUG-2000; 2000US-0225447P.

PR 14-AUG-2000; 2000US-0225757P.

PR 14-AUG-2000; 2000US-0225758P.

PR 18-AUG-2000; 2000US-0225759P.

PR 22-AUG-2000; 2000US-0226681P.

PR 22-AUG-2000; 2000US-0226686P.

PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 05-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.

```
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 06-DEC-2000; 2000US-0256719P.
PR 08-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
PI WPI; 2001-483426/52.
XX
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and metastasis.
PT
XX Disclosure; SEQ ID NO 38570; 3071pp + Sequence Listing; English.
XX
XX AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
CC amino acid sequences given in AAK91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting the
CC nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/hematopoietic-related diseases, especially
CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/hematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
CC represent sequences used in the exemplification of the present invention
XX
XX Sequence 25576 BP; 7153 A; 6020 C; 6060 G; 6343 T; 0 U; 0 Other;
SQ
Query Match 83.2%; Score 20.8; DB 4; Length 25576;
Best Local Similarity 91.7%; Pred. No. 44;
Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 AAAAAAAAAATCAGGCGAGGCGATG 24
Db 12225 AAAAAAAAAATCAGGCGAGGCGATG 12248
RESULT 18
AAL05618
```

```
ID AAL05618 standard; DNA; 25576 BP.
XX
XX AAL05618;
AC
XX
XX 21-NOV-2001 (first entry)
DT
XX
XX Human reproductive system related antigen DNA SEQ ID NO: 8306.
DE
XX
XX Human; reproductive system related antigen; reproductive system disorder;
KW cancer; gene therapy; ds.
KW
XX
XX Homo sapiens.
OS
XX
XX WO200155320-A2.
PN
XX
XX 02-AUG-2001.
PD
XX
XX 17-JAN-2001; 2001WO-US001339.
PF
XX
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216800P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
```

PR 14-SEP-2000; 2000US-0232399P.
 PR 14-SEP-2000; 2000US-0232400P.
 PR 14-SEP-2000; 2000US-0232401P.
 PR 14-SEP-2000; 2000US-0233663P.
 PR 14-SEP-2000; 2000US-0233664P.
 PR 14-SEP-2000; 2000US-0233665P.
 PR 21-SEP-2000; 2000US-0234223P.
 PR 21-SEP-2000; 2000US-0234274P.
 PR 25-SEP-2000; 2000US-0234997P.
 PR 25-SEP-2000; 2000US-0234998P.
 PR 26-SEP-2000; 2000US-0235484P.
 PR 27-SEP-2000; 2000US-0235834P.
 PR 27-SEP-2000; 2000US-0235836P.
 PR 29-SEP-2000; 2000US-0236327P.
 PR 29-SEP-2000; 2000US-0236367P.
 PR 29-SEP-2000; 2000US-0236368P.
 PR 29-SEP-2000; 2000US-0236369P.
 PR 29-SEP-2000; 2000US-0236370P.
 PR 02-OCT-2000; 2000US-0236802P.
 PR 02-OCT-2000; 2000US-0237037P.
 PR 02-OCT-2000; 2000US-0237038P.
 PR 02-OCT-2000; 2000US-0237039P.
 PR 02-OCT-2000; 2000US-0237040P.
 PR 13-OCT-2000; 2000US-0239935P.
 PR 13-OCT-2000; 2000US-0239937P.
 PR 20-OCT-2000; 2000US-0240960P.
 PR 20-OCT-2000; 2000US-0241221P.
 PR 20-OCT-2000; 2000US-0241785P.
 PR 20-OCT-2000; 2000US-0241786P.
 PR 20-OCT-2000; 2000US-0241787P.
 PR 20-OCT-2000; 2000US-0241808P.
 PR 20-OCT-2000; 2000US-0241809P.
 PR 20-OCT-2000; 2000US-0241826P.
 PR 01-NOV-2000; 2000US-0244617P.
 PR 08-NOV-2000; 2000US-0246474P.
 PR 08-NOV-2000; 2000US-0246475P.
 PR 08-NOV-2000; 2000US-0246476P.
 PR 08-NOV-2000; 2000US-0246477P.
 PR 08-NOV-2000; 2000US-0246478P.
 PR 08-NOV-2000; 2000US-0246523P.
 PR 08-NOV-2000; 2000US-0246524P.
 PR 08-NOV-2000; 2000US-0246525P.
 PR 08-NOV-2000; 2000US-0246526P.
 PR 08-NOV-2000; 2000US-0246527P.
 PR 08-NOV-2000; 2000US-0246528P.
 PR 08-NOV-2000; 2000US-0246532P.
 PR 08-NOV-2000; 2000US-0246609P.
 PR 08-NOV-2000; 2000US-0246610P.
 PR 08-NOV-2000; 2000US-0246611P.
 PR 08-NOV-2000; 2000US-0246613P.
 PR 17-NOV-2000; 2000US-0249207P.
 PR 17-NOV-2000; 2000US-0249208P.
 PR 17-NOV-2000; 2000US-0249209P.
 PR 17-NOV-2000; 2000US-0249210P.
 PR 17-NOV-2000; 2000US-0249211P.
 PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249244P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249246P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249266P.
 PR 17-NOV-2000; 2000US-0249267P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.

PR 05-DEC-2000; 2000US-0256719P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 08-DEC-2000; 2000US-0251990P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Rosen CA, Barash SC, Ruben SM;
 XX
 XX WPI; 2001-465570/50.
 XX
 XX Isolated nucleic acid molecule encoding a reproductive system antigen is
 PT used in preventing, treating or ameliorating a medical condition.
 XX
 XX Disclosure; SEQ ID NO 8306; 1297pp + Sequence Listing; English.
 XX
 XX The present invention provides the protein and coding sequences of a
 CC number of human reproductive system related antigens. These can be used
 CC in the prevention and treatment of reproductive system disorders,
 CC including cancer. The present sequence is a genomic sequence encoding a
 CC protein of the invention
 XX
 XX Sequence 25576 BP; 7153 A; 6020 C; 6060 G; 6343 T; 0 U; 0 Other;
 SQ
 Query Match 83.2%; Score 20.8; DB 4; Length 25576;
 Best Local Similarity 91.7%; Pred. No. 44;
 Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 AAAAAAAAAATCAGGCCAGGCATG 24
 DB 12225 AAAAAAAAAATCAGGCCAGGCATG 12248
 RESULT 19
 ACN44934_1/c
 Continuation (2 of 4) of ACN44934 from base 100001 (Human genomic sequence hCG1811271.
 WP Sequence split into 4 fragments LOCUS ACN44934 Accession ACN44934
 WP Fragment Name Begin End
 WP ACN44934_0 1 110000
 WP ACN44934_1 100001 210000
 WP ACN44934_2 200001 310000
 WP ACN44934_3 300001 401616
 Query Match 83.2%; Score 20.8; DB 11; Length 110000;
 Best Local Similarity 91.7%; Pred. No. 57;
 Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 2 AAAAAAAAAATCAGGCCAGGCATG 25
 DB 40060 AAAAAAAAAATCAGGCCAGGCATG 40037
 RESULT 20
 ADZ46976_3/c
 Continuation (4 of 6) of ADZ46976 from base 300001 (Human phosphodiesterase 4B DNA SEQ
 WP Sequence split into 6 fragments LOCUS ADZ46976 Accession ADZ46976
 WP Fragment Name Begin End
 WP ADZ46976_0 1 110000
 WP ADZ46976_1 100001 210000
 WP ADZ46976_2 300001 310000
 WP ADZ46976_3 300001 410000
 WP ADZ46976_4 400001 510000
 WP ADZ46976_5 500001 599001
 Query Match 83.2%; Score 20.8; DB 14; Length 110000;
 Best Local Similarity 91.7%; Pred. No. 57;
 Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

XX	Human secreted protein 5' EST, SEQ ID NO: 21508.
DE	
XX	
XX	Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation; gene therapy; chromosome mapping; ss.
KW	
XX	
XX	Homo sapiens.
OS	
XX	EP1033401-A2.
PN	
XX	
XX	06-SEP-2000.
PD	
XX	
XX	21-FEB-2000; 2000EP-00200610.
PF	
XX	
XX	26-FEB-1999; 99US-0122487P.
PR	
XX	
XX	(GEST) GENSET.
PA	
XX	
XX	Dumas Milne Edwards J, Duclert A, Giordano J;
PI	
XX	
XX	WPI; 2000-500381/45.
DR	
XX	
XX	New nucleic acid that is a 5' expressed sequence tag (5' EST) for obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for diagnostic, forensic, gene therapy and chromosome mapping procedures.
PT	
PT	
XX	
XX	Claim 1; SEQ ID NO 21508; 71pp + Sequence Listing; English.
PS	
XX	
XX	The present sequence is one of a large number of 5' ESTs derived from mRNAs encoding secreted proteins. No ORF has yet been conclusively identified within the present sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs derived from 30 different tissues. EST sequences usually correspond mainly to the 3' untranslated region (UTR) of the mRNA because they are often obtained from oligo-dT primed cDNA libraries. Such ESTs are not well suited for isolating cDNA sequences derived from the 5' ends of mRNAs and even in those cases where longer cDNA sequences have been obtained, the full 5' UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and chromosome mapping procedures. They are used to obtain upstream regulatory sequences and to design expression and secretion vectors
CC	
XX	
SQ	Sequence 138 BP; 25 A; 37 C; 30 G; 46 T; 0 U; 0 Other;
	Query Match 80.8%; Score 20.2; DB 3; Length 138;
	Best Local Similarity 88.0%; Pred No. 34;
	Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy	1 AAAAAAAAAATCACGGCCAGGCATGG 25
Db	131 AAAAAAAAAAAAGGCCAGGCATGG 107
RESULT 23	
AAC14155/c	
ID	AAC14155 standard; cDNA; 206 BP.
XX	
XX	AAC14155;
AC	
XX	
XX	06-OCT-2000 (first entry)
DT	
XX	
XX	Human secreted protein 5' EST, SEQ ID NO: 18230.
DE	
XX	
XX	Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation; gene therapy; chromosome mapping; ss.
KW	
XX	
XX	Homo sapiens.
OS	
XX	
XX	EP1033401-A2.
PN	
XX	
XX	06-SEP-2000.
PD	
XX	

```
PF 21-FEB-2000; 2000EP-00200610.
XX
PR 26-FEB-1999; 99US-0122487P.
XX
PA (GEST ) GENSET.
XX
PI Dumas Milne Edwards J, Duclert A, Giordano J;
XX
PI WPI; 2000-500381/45.
XX
DR
XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures.
XX
XX Claim 1; SEQ ID NO 18230; 71pp + Sequence Listing; English.
XX
XX The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC identified within the present sequence. The 5' ESTs were prepared from
CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
CC sequences usually correspond mainly to the 3' untranslated region (UTR)
CC of the mRNA because they are often obtained from oligo-dT primed cDNA
CC libraries. Such ESTs are not well suited for isolating cDNA sequences
CC derived from the 5' ends of mRNAs and even in those cases where longer
CC cDNA sequences have been obtained, the full 5' UTR is rarely included. 5'
CC ESTs are derived from mRNAs with intact 5' ends and can therefore be used
CC to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in
CC diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC They are used to obtain upstream regulatory sequences and to design
CC expression and secretion vectors
XX
XX Sequence 206 BP; 47 A; 36 C; 42 G; 76 T; 0 U; 5 Other;
SQ
Query Match 80.8%; Score 20.2; DB 3; Length 206;
Best Local Similarity 88.0%; Pred. No. 36;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 AAAAAAAAAATCAGGCCAGCATGG 25
Db 201 AAAAAAAAAATCAGGCCAGCATGG 177
RESULT 24
ACL56670/c
ID ACL56670 standard; cDNA; 302 BP.
XX
XX ACL56670;
XX
XX 24-MAR-2005 (first entry)
XX
XX Human colon cancer differentially expressed polynucleotide, SEQ ID:2805.
XX
XX Differential expression; diagnosis; therapy; drug screening; cancer;
XX neoplasm; colon tumor; breast tumor; pancreas tumor; cytostatic; vaccine;
XX ss.
XX
XX Homo sapiens.
XX
XX WO2005000087-A2.
XX
XX 06-JAN-2005.
XX
XX 13-MAY-2004; 2004WO-US015421.
XX
XX 03-JUN-2003; 2003US-0475872P.
XX
XX (CHIR ) CHIRON CORP.
XX
XX Randazzo F, Moler E, Escobedo J, Garcia PD;
PI
XX WPI; 2005-075421/08.
XX
XX New isolated polynucleotides, which are differentially expressed in colon
PT
```

```
PT cancer cell, useful for treating cancer, e.g. colon cancer, breast
XX cancer, or pancreatic cancer.
XX
XX Claim 1; SEQ ID NO 2805; 97pp; English.
XX
XX The invention relates to 9672 polynucleotides (ACL53866-ACL63537) which
CC are differentially expressed in colon cancer cells. The invention also
CC relates to vectors and host cells comprising a differentially expressed
CC polynucleotide of the invention; a method for detecting a cancerous cell
CC by detection of a gene product of the polynucleotides; a method for
CC inhibiting a cancerous phenotype of a cell by inhibiting a gene product
CC of the polynucleotides; a method of treating an individual with cancer by
CC administration of a modulator of a gene product of the polynucleotides;
CC and an isolated antibody that specifically binds to a polypeptide encoded
CC by one of the 9672 polynucleotides. The polynucleotides, polypeptides,
CC antibodies, and methods are useful for the detection of cancerous cells;
CC for the diagnosis, prognosis and management of cancer; for the
CC identification of agents that modulate the phenotype of cancerous cells;
CC for the identification of therapeutic targets for cancer chemotherapy;
CC and for the treatment of cancer, especially colon cancer and metastasized
CC colon cancer, but also breast or pancreatic cancer. The polynucleotides
CC are also useful as a source of probes or primers for use in diagnostic
CC methods. The differentially expressed polynucleotides or their encoded
CC proteins can additionally be used as vaccines to modulate primary immune
CC responses for the prevention or treatment of cancer. The present sequence
CC represents a specifically claimed polynucleotide which is differentially
CC expressed in colon cancer. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 302 BP; 76 A; 61 C; 61 G; 104 T; 0 U; 0 Other;
SQ
Query Match 80.8%; Score 20.2; DB 14; Length 302;
Best Local Similarity 88.0%; Pred. No. 39;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 AAAAAAAAAATCAGGCCAGCATGG 25
Db 139 AAAAAAAAAAAGGCCAGCATGG 115
RESULT 25
AAK74375
ID AAK74375 standard; DNA; 339 BP.
XX
XX AAK74375;
XX
XX 07-NOV-2001 (first entry)
XX
XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:29187.
XX
XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
XX cytostatic; gene therapy; vaccine; metastasis; ds.
XX
XX Homo sapiens.
XX
XX WO200157182-A2.
XX
XX 09-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US001354.
XX
XX 31-JAN-2000; 2000US-0179065P.
XX
XX 04-FEB-2000; 2000US-0180628P.
XX
XX 24-FEB-2000; 2000US-0184664P.
XX
XX 02-MAR-2000; 2000US-0186350P.
XX
XX 16-MAR-2000; 2000US-0189874P.
XX
XX 17-MAR-2000; 2000US-0190076P.
XX
XX 18-APR-2000; 2000US-0198123P.
XX
XX 19-MAY-2000; 2000US-0205515P.
XX
XX 07-JUN-2000; 2000US-0209467P.
XX
XX 28-JUN-2000; 2000US-0214886P.
```

PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226682P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 03-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.

PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-483426/52.

Nucleic acids encoding human immune/hematopoietic antigen polypeptides, useful for preventing, diagnosing and/or treating cancers and metastasis.

Disclosure; SEQ ID NO 29187; 3071pp + Sequence Listing; English.

AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I) amino acid sequences given in AAK62170 to AAK61921. (I) have cytostatic activity, and can be used in gene therapy and vaccine production. (I) proteins, and polynucleotides may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate (I) expression. For example, they may be used to treat disorders associated with decreased expression by rectifying mutations or deletions in a patient's genome that affect the activity of (I) by expressing inactive proteins or to

CC supplement the patients own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I), by inserting the
 CC nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/haematopoietic-related diseases, especially
 CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
 CC to AAK87694 represent human immune/haematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
 CC represent sequences used in the exemplification of the present invention
 CC
 SQ Sequence 339 BP; 126 A; 50 C; 82 G; 81 T; 0 U; 0 Other;

Query Match 80.8%; Score 20.2; DB 4; Length 339;
 Best Local Similarity 88.0%; Pred. No. 39;
 Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAAAAAATCAGCGCCAGGCATGG 25
 ||||| ||||| ||||| ||||| |||||
 Db 234 AAAAAAAAAACAAAGCGCCAGGCATGG 258

RESULT 26
 AAS37660
 ID AAS37660 standard; cDNA; 344 BP.

AC AAS37660;

XX 17-DEC-2001 (first entry)

DE Novel human diagnostic and therapeutic gene #718.

XX Human; cancer; breast; lung; colon; prostate; cytostatic; diagnostic; ss.

OS Homo sapiens.

XX WO200166753-A2.

PN 13-SEP-2001.

XX 09-MAR-2001; 2001WO-US0007787.

XX 09-MAR-2000; 2000US-0188609P.

XX (CHIR) CHIRON CORP.

PA (HYSE-) HYSEQ INC.

XX Williams LT, Escobedo J, Innis MA, Garcia PD, Sudduth-Klinger J;

PI Reinhard C, Randazzo F, Kennedy GC, Pot D, Kassam A, Lamson G;

PI Drmanac R, Crkvenjakov R, Dickson M, Drmanac S, Labat I;

PI Leshkowitz D, Kita D, Garcia V, Jones WL, Stache-Crain B;

XX WPI; 2001-530177/58.

XX New polynucleotides and polypeptides, useful for diagnosis and treatment
 of breast, lung and colon cancer.

XX Claim 1; Page 775-776; 1193pp; English.

XX The invention relates to new polynucleotides and polypeptides, useful for
 diagnosis and treatment of breast, lung and colon cancer. The sequences
 CC can be used in detecting differentially expressed genes correlated with a
 CC cancerous state of a mammalian cell, comprising detecting at least one
 CC differentially expressed gene product in a test sample derived from a
 CC cell suspected of being cancerous. They can also be used to inhibit
 CC tumour growth by modulating expression of a gene product. AAS36943-
 CC AAS39338 represent novel human diagnostic and therapeutic coding
 CC sequences of the invention

XX Sequence 344 BP; 125 A; 62 C; 61 G; 96 T; 0 U; 0 Other;

Query Match 80.8%; Score 20.2; DB 4; Length 344;
 Best Local Similarity 88.0%; Pred. No. 40;
 Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAAAAAATCAGCGCCAGGCATGG 25
 ||||| ||||| ||||| ||||| |||||
 Db 237 AAAAAAAAAATCTTGGCCAGGCATGG 261

RESULT 27

ABK95389/C

ID ABK95389 standard; DNA; 384 BP.

XX ABK95389;

XX 24-SEP-2002 (first entry)

XX Human retina specific gene MPP4 sequence containing exon 13.

XX Human; ds; MPP4; C7orf9; C12orf7; F379; retina specific gene; AMD;

XX age-related macular degeneration; blindness; gene therapy;

XX Ophthalmological; transgenic.

OS Homo sapiens.

XX WO200244366-A2.

XX 06-JUN-2002.

XX 29-NOV-2001; 2001WO-EP013940.

XX 29-NOV-2000; 2000US-0253751P.

XX (MULT-) MULTIGENE BIOTECH GMBH.

XX Stoehr HB, Weber BHF;

XX WPI; 2002-508512/54.

XX Novel nucleic acids encoding retina-specific human protein C7orf9,
 C12orf7, MPP4 or F379, useful for diagnosing age-related macular
 PT degeneration or predisposition for macular degeneration, and in gene
 PT therapy techniques.

XX Claim 1; Fig 6c; 120pp; English.

XX The invention relates to isolated nucleic acid encoding retina-specific
 CC human protein C7orf9, C12orf7, MPP4 or F379 or a fragment, derivative or
 CC allelic variation of the above mentioned nucleic acid sequences. Also
 CC included are a recombinant vector containing the nucleic acid, a
 CC recombinant host cell which contains the vector and expresses the
 CC protein, an inhibitor characterised in that it can suppress the activity
 CC of the protein, treating macular degeneration or a predisposition for
 CC macular degeneration, comprising administering to a mammalian subject a
 CC reagent which decreases, inhibits or increases expression of C7orf9,
 CC C12orf7, MPP4 and/or F379 or which leads to the expression of a
 CC biologically active C7orf9, C12orf7, MPP4 and/or F379 protein and a
 CC transgenic non-human animal comprising at least one of the nucleic acids
 CC (active or inactivated). The nucleic acid or protein is useful for
 CC diagnosing macular degeneration, preferably age-related macular
 CC degeneration (AMD) or a predisposition for macular degeneration. The
 CC reagent used in the diagnosis is a C7orf9-, C12orf7-, MPP4- or F379 -
 CC specific nucleic acid probe, or anti-C7orf9, anti-C12orf7, anti-MPP4 or
 CC anti-F379-antibody. The reagent is detectably labeled, with a compound
 CC such as a radioisotope, a bioluminescent compound, a chemiluminescent
 CC compound, a fluorescent compound, a metal chelate or an enzyme. Fragments
 CC of the nucleic acid are useful as probes or primers in a diagnostic
 CC assay, and for identifying further factors involved in development and
 CC progression of macular degeneration. The proteins encoded by the nucleic
 CC acid are useful to identify further unrelated proteins which are
 CC associated with macular degeneration and for use in screening methods
 CC based on protein/protein interactions. The nucleic acid is also useful as
 CC reagents for detecting differences between normal and aberrant expression
 CC of the protein. the nucleic acid is also useful in gene therapy
 CC techniques, and can be used for gene targeting and/or gene replacement
 CC for restoring a mutant gene or for creating a mutant gene via homologous

CC recombination. the protein can be used to identify other proteins
CC involved in development and progression of macular degeneration. The
CC present sequence is a genomic DNA fragment containing an exon of a retina
CC specific gene of the invention

XX
SQ Sequence 384 BP; 95 A; 85 C; 86 G; 118 T; 0 U; 0 Other;

Query Match 80.8%; Score 20.2; DB 6; Length 384;

Best Local Similarity 88.0%; Pred. No. 40;

Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAAAAAATCAGCCAGGCATGG 25

DB 125 AAGAAATTCAGCCAGGCATGG 101

RESULT 28

AAK83882/C

ID AAK83882 standard; DNA; 392 BP.

XX AC AAK83882;

XX DT 07-NOV-2001 (first entry)

XX DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:38694.

XX KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;

XX KW cytostatic; gene therapy; vaccine; metastasis; ds.

XX OS Homo sapiens.

XX PN WO200157182-A2.

XX PD 09-AUG-2001.

XX PF 17-JAN-2001; 2001WO-US001354.

XX PR 31-JAN-2000; 2000US-0179065P.

PR 04-FEB-2000; 2000US-0180628P.

PR 24-FEB-2000; 2000US-0184664P.

PR 02-MAR-2000; 2000US-0186350P.

PR 16-MAR-2000; 2000US-0189874P.

PR 17-MAR-2000; 2000US-0190076P.

PR 18-APR-2000; 2000US-0198123P.

PR 19-MAY-2000; 2000US-0205515P.

PR 07-JUN-2000; 2000US-0209467P.

PR 28-JUN-2000; 2000US-0214886P.

PR 30-JUN-2000; 2000US-0215135P.

PR 07-JUL-2000; 2000US-0216647P.

PR 07-JUL-2000; 2000US-0216880P.

PR 11-JUL-2000; 2000US-0217487P.

PR 11-JUL-2000; 2000US-0217496P.

PR 14-JUL-2000; 2000US-0218290P.

PR 26-JUL-2000; 2000US-0220363P.

PR 26-JUL-2000; 2000US-0220364P.

PR 14-AUG-2000; 2000US-0224519P.

PR 14-AUG-2000; 2000US-0224519P.

PR 14-AUG-2000; 2000US-0225213P.

PR 14-AUG-2000; 2000US-0225214P.

PR 14-AUG-2000; 2000US-0225266P.

PR 14-AUG-2000; 2000US-0225267P.

PR 14-AUG-2000; 2000US-0225268P.

PR 14-AUG-2000; 2000US-0225270P.

PR 14-AUG-2000; 2000US-0225447P.

PR 14-AUG-2000; 2000US-0225757P.

PR 14-AUG-2000; 2000US-0225758P.

PR 14-AUG-2000; 2000US-0225759P.

PR 22-AUG-2000; 2000US-0226279P.

PR 22-AUG-2000; 2000US-0226681P.

PR 22-AUG-2000; 2000US-0226686P.

PR 23-AUG-2000; 2000US-0227182P.

PR 23-AUG-2000; 2000US-0227009P.

PR 30-AUG-2000; 2000US-0228924P.

PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232377P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 13-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.

```
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249246P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 08-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254057P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-483426/52.
XX
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
XX useful for preventing, diagnosing and/or treating cancers and metastasis.
XX
XX Disclosure; SEQ ID NO 38694; 3071pp + Sequence Listing; English.
XX
XX AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
XX amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
XX activity, and can be used in gene therapy and vaccine production. (I)
XX proteins and polynucleotides may be used in the prevention, diagnosis and
XX treatment of diseases associated with inappropriate (I) expression. For
XX example, they may be used to treat disorders associated with decreased
XX expression by rectifying mutations or deletions in a patient's genome
XX that affect the activity of (I) by expressing inactive proteins or to
XX supplement the patients own production of (I). Additionally, (I)
XX polynucleotides may be used to produce the secreted (I), by inserting the
XX nucleic acids into a host cell and culturing the cell to express the
XX protein. (I) proteins and polynucleotides may be used to prevent,
XX diagnose and treat immune/haematopoietic-related diseases, especially
XX cancers and cancer metastases of haematopoietic-derived cells. AAK64703
XX to AAK87694 represent human immune/haematopoietic antigen genomic
XX sequences from the present invention. AAK54942 to AAK54950 and AAM82169
XX represent sequences used in the exemplification of the present invention
XX
XX Sequence 392 BP; 100 A; 79 C; 79 G; 134 T; 0 U; 0 Other;
XX
XX Query Match 80.8%; Score 20.2; DB 4; Length 392;
XX Best Local Similarity 88.0%; Pred. No. 40;
XX Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
XX
XX Qy 1 AAAAAAAAAATCAGCGCCAGGCATGG 25
XX ||||| ||||| ||||| |||||
XX Db 208 AAAAAAAAAAAAGGCCAGGCATGG 184
XX
XX RESULT 29
XX AAI87906/c
XX ID AAI87906 standard; cDNA; 403 BP.
XX
XX AC AAI87906;
XX
XX Sequence 403 BP; 86 A; 120 C; 95 G; 102 T; 0 U; 0 Other;
XX
XX Query Match 80.8%; Score 20.2; DB 4; Length 403;
XX Best Local Similarity 88.0%; Pred. No. 41;
XX Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
XX
XX Qy 1 AAAAAAAAAATCAGCGCCAGGCATGG 25
XX ||||| ||||| ||||| |||||
XX Db 94 AAAAAAAAAAAAGGCCAGGCATGG 70
XX
XX RESULT 30
XX ABL82000/c
XX ID ABL82000 standard; cDNA; 465 BP.
XX
XX AC ABL82000;
XX
XX Qy 17-MAY-2002 (first entry)
XX
XX Human ovarian cancer related cDNA clone SEQ ID NO:4978.
XX
XX Human; ovarian cancer; ovarian tumour; cytostatic; gene; ss.
XX
XX Homo sapiens.
XX
XX WO200192581-A2.
XX
XX PN
```

```
XX
XX 06-NOV-2001 (first entry)
XX
XX Human polynucleotide SEQ ID NO 7966.
XX
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
XX vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
XX tissue growth factor; immunomodulatory; cancer; leukaemia;
XX nervous system disorders; arthritis; inflammation; ss.
XX
XX Homo sapiens.
XX
XX WO200164835-A2.
XX
XX 07-SEP-2001.
XX
XX 26-FEB-2001; 2001WO-US004927.
XX
XX 28-FEB-2000; 2000US-00515126.
XX
XX 18-MAY-2000; 2000US-00577409.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Drmanac RT;
XX
XX WPI; 2001-514838/56.
XX
XX P-PSDB; AAO07975.
XX
XX Isolated nucleic acids and polypeptides, useful for preventing diagnosing
XX and treating e.g. leukemia, inflammation and immune disorders.
XX
XX Claim 1; SEQ ID NO 7966; 1399pp + Sequence Listing; English.
XX
XX The invention relates to human polynucleotides (AAI79941-AAI93841) and
XX the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to
XX cytokine, cell proliferation or cell differentiation or which may induce
XX production of other cytokines in other cell populations. The
XX polynucleotides and polypeptides are useful in gene therapy, vaccines or
XX peptide therapy. The polypeptides have various cytokine-like activities,
XX e.g. stem cell growth factor activity, haematopoiesis regulating
XX activity, tissue growth factor activity, immunomodulatory activity and
XX activin/inhibin activity and may be useful in the diagnosis and/or
XX treatment of cancer, leukaemia, nervous system disorders, arthritis and
XX inflammation. Note: The sequence data for this patent did not form part
XX of the printed specification, but was obtained in electronic format
XX directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 403 BP; 86 A; 120 C; 95 G; 102 T; 0 U; 0 Other;
XX
XX Query Match 80.8%; Score 20.2; DB 4; Length 403;
XX Best Local Similarity 88.0%; Pred. No. 41;
XX Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
XX
XX Qy 1 AAAAAAAAAATCAGCGCCAGGCATGG 25
XX ||||| ||||| ||||| |||||
XX Db 94 AAAAAAAAAAAAGGCCAGGCATGG 70
XX
XX RESULT 30
XX ABL82000/c
XX ID ABL82000 standard; cDNA; 465 BP.
XX
XX AC ABL82000;
XX
XX Qy 17-MAY-2002 (first entry)
XX
XX Human ovarian cancer related cDNA clone SEQ ID NO:4978.
XX
XX Human; ovarian cancer; ovarian tumour; cytostatic; gene; ss.
XX
XX Homo sapiens.
XX
XX WO200192581-A2.
XX
XX PN
```


(HYSE-) HYSEQ INC.
PA Escobedo J, Garcia PD, Sudduth-Klinger J, Reinhard C, Randazzo F;
XX Lamson G, Scott EM, Zhang G, Kaasam A, Pot D, Labat I;
PI WPI; 2002-241905/29.
XX
DR
XX New nucleic acid for producing a polypeptide, detecting differentially
PT expressed genes correlated with a cancerous state of a mammalian cell,
PT and inhibiting tumor growth.
XX
XX Claim 1; SEQ ID NO 1717; 883pp + Sequence Listing; English.
PS
XX The invention relates to an isolated polynucleotide (ABN27253-ABN33262)
CC with cytostatic activity. The polynucleotide is used to produce a
CC polypeptide, to detect differentially expressed genes correlated with a
CC cancerous state of a mammalian cell and to inhibit tumor growth. The
CC polynucleotide is used as a probe in mapping and tissue profiling. The
CC encoded polypeptide and antibodies to the polypeptide can also be used
CC for therapeutic and diagnostic purposes. The polynucleotide is useful for
CC gene therapy. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 480 BP; 158 A; 97 C; 132 G; 93 T; 0 U; 0 Other;
SQ
Query Match 80.8%; Score 20.2; DB 6; Length 480;
Best Local Similarity 88.0%; Pred. No. 42;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 AAAAAAAAAATCAGCGCCAGGCATGG 25
Db 238 AAAAAAAAAATCAAGCGCGGCATGG 262
RESULT 33
ACL60088
ID ID ACL60088 standard; cDNA; 480 BP.
XX
AC ACL60088;
XX
XX 24-MAR-2005 (first entry)
DT
XX Human colon cancer differentially expressed polynucleotide, SEQ ID:6223.
DE
XX Differential expression; diagnosis; therapy; drug screening; cancer;
KW neoplasm; colon tumor; breast tumor; pancreas tumor; cytostatic; vaccine;
KW ss.
KW
XX Homo sapiens.
OS
XX WO200500087-A2.
PN
XX 06-JAN-2005.
PD
XX 13-MAY-2004; 2004WO-US015421.
XX
XX 03-JUN-2003; 2003US-0475872P.
XX
XX (CHIR) CHIRON CORP.
PA
XX Randazzo F, Moler E, Escobedo J, Garcia PD;
PI WPI; 2005-075421/08.
XX
XX New isolated polynucleotides, which are differentially expressed in colon
XX cancer cell, useful for treating cancer, e.g. colon cancer, breast
XX cancer, or pancreatic cancer.
XX
XX Claim 1; SEQ ID NO 6223; 97pp; English.
PS
XX The invention relates to 9672 polynucleotides (ACL53866-ACL63537) which
CC are differentially expressed in colon cancer cells. The invention also

relates to vectors and host cells comprising a differentially expressed
polynucleotide of the invention; a method for detecting a cancerous cell
by detection of a gene product of the polynucleotides; a method for
inhibiting a cancerous phenotype of a cell by inhibiting a gene product
of the polynucleotides; a method of treating an individual with cancer by
administration of a modulator of a gene product of the polynucleotides;
and an isolated antibody that specifically binds to a polypeptide encoded
by one of the 9672 polynucleotides. The polynucleotides, polypeptides,
antibodies, and methods are useful for the detection of cancerous cells;
for the diagnosis, prognosis and management of cancer; for the
identification of agents that modulate the phenotype of cancerous cells;
and for the treatment of cancer, especially colon cancer and metastasized
colon cancer, but also breast or pancreatic cancer. The polynucleotides
are also useful as a source of probes or primers for use in diagnostic
methods. The differentially expressed polynucleotides or their encoded
proteins can additionally be used as vaccines to modulate primary immune
responses for the prevention or treatment of cancer. The present sequence
represents a specifically claimed polynucleotide which is differentially
expressed in colon cancer. Note: The sequence data for this patent did
not form part of the printed specification, but was obtained in
electronic format directly from WIPO at
ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 480 BP; 158 A; 97 C; 132 G; 93 T; 0 U; 0 Other;
SQ
Query Match 80.8%; Score 20.2; DB 14; Length 480;
Best Local Similarity 88.0%; Pred. No. 42;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 AAAAAAAAAATCAGCGCCAGGCATGG 25
Db 238 AAAAAAAAAATCAAGCGCGGCATGG 262
RESULT 34
ABA59354
ID ID ABA59354 standard; DNA; 484 BP.
XX
AC ABA59354;
XX
XX 01-FEB-2002 (first entry)
DT
XX Human foetal liver single exon nucleic acid probe #7659.
DE
XX Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
KW
XX Homo sapiens.
OS
XX WO200157277-A2.
PN
XX 09-AUG-2001.
PD
XX 30-JAN-2001; 2001WO-US000669.
XX
XX 04-FEB-2000; 2000US-0180312P.
PR
XX 26-MAY-2000; 2000US-0207456P.
PR
XX 30-JUN-2000; 2000US-00608408.
PR
XX 03-AUG-2000; 2000US-00632366.
PR
XX 21-SEP-2000; 2000US-0234687P.
PR
XX 27-SEP-2000; 2000US-0236359P.
PR
XX 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLR-) MOLECULAR DYNAMICS INC.
PA
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI WPI; 2001-483447/52.
XX
XX Human genome-derived single exon nucleic acid probes useful for analyzing
XX gene expression in human fetal liver.
XX
XX Claim 1; SEQ ID NO 7659; 639pp + Sequence Listing; English.
PS

XX The invention relates to a single exon nucleic acid probe for measuring
CC human gene expression in a sample derived from human foetal liver. The
CC single exon nucleic acid probes may be used for predicting, measuring and
CC displaying gene expression in samples derived from human fetal liver. The
CC present sequence is a single exon nucleic acid probe of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 484 BP; 166 A; 95 C; 138 G; 85 T; 0 U; 0 Other;

Query Match 80.8%; Score 20.2; DB 4; Length 484;
Best Local Similarity 88.0%; Pred. No. 42;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAAAAAATCAGCGCCAGGCATGG 25
Db 55 AAAAAACAAGCAAGCGCCAGGCATGG 79

RESULT 35
AAI39155
ID AAI39155 standard; DNA; 484 BP.
XX
AC AAI39155;
XX
DT 17-OCT-2001 (first entry)
XX
DE Probe #7841 used to measure gene expression in human placenta sample.
XX
XX Probe; microarray; human; placenta; antenatal diagnosis;
KW genetic disorder; ss.
XX
OS Homo sapiens.
XX
XX
PN WO200157272-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000663.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PN WO200157272-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000663.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI WPI; 2001-488897/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human placenta.
XX
PS Claim 25; SEQ ID NO 7841; 654pp; English.
XX
XX The present invention relates to single exon nucleic acid probes (SENPs).
CC The present sequence is one such probe. The probes are useful for
CC producing a microarray for predicting, measuring and displaying gene
CC expression in samples derived from human placenta. The probes are useful
CC for antenatal diagnosis of human genetic disorders
XX
SQ Sequence 484 BP; 166 A; 95 C; 138 G; 85 T; 0 U; 0 Other;

Query Match 80.8%; Score 20.2; DB 4; Length 484;
Best Local Similarity 88.0%; Pred. No. 42;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAAAAAATCAGCGCCAGGCATGG 25
Db 55 AAAAAACAAGCAAGCGCCAGGCATGG 79

RESULT 37
AAK07570
ID AAK07570 standard; DNA; 484 BP.
XX
AC AAK07570;
XX
DT 05-NOV-2001 (first entry)
XX
XX Human brain expressed single exon probe SEQ ID NO: 7561.
XX

QY 1 AAAAAAAAAATCAGCGCCAGGCATGG 25
Db 55 AAAAAACAAGCAAGCGCCAGGCATGG 79

RESULT 36
AAK33368
ID AAK33368 standard; DNA; 484 BP.
XX
AC AAK33368;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human bone marrow expressed single exon probe SEQ ID NO: 7925.
XX
XX Human; bone marrow expressed exon; gene expression analysis; probe;
KW microarray; cancer; leukaemia; lymphoma; myeloma; ss.
XX
OS Homo sapiens.
XX
XX WO200157276-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000668.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI WPI; 2001-488900/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human bone marrow.
XX
PS Example 4; SEQ ID NO 7925; 658pp + Sequence Listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukaemia and myeloma. The present sequence is one of
CC the probes of the invention
XX
SQ Sequence 484 BP; 166 A; 95 C; 138 G; 85 T; 0 U; 0 Other;

Query Match 80.8%; Score 20.2; DB 4; Length 484;
Best Local Similarity 88.0%; Pred. No. 42;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAAAAAATCAGCGCCAGGCATGG 25
Db 55 AAAAAACAAGCAAGCGCCAGGCATGG 79

RESULT 37
AAK07570
ID AAK07570 standard; DNA; 484 BP.
XX
AC AAK07570;
XX
DT 05-NOV-2001 (first entry)
XX
XX Human brain expressed single exon probe SEQ ID NO: 7561.
XX

KW Human; brain expressed exon; gene expression analysis; probe; microarray;
KW Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer;
KW ss.

XX Homo sapiens.

OS WO200157275-A2.

PN 09-AUG-2001.

PD 30-JAN-2001; 2001WO-US000667.

XX 04-FEB-2000; 2000US-0180312P.

XX 26-MAY-2000; 2000US-0207456P.

PR 30-JUN-2000; 2000US-00608408.

PR 03-AUG-2000; 2000US-00632366.

PR 21-SEP-2000; 2000US-0234687P.

PR 27-SEP-2000; 2000US-0236359P.

PR 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

PA Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-483446/52.

XX Single exon nucleic acid probes for analyzing gene expression in human

XX brains.

XX Example 4; SEQ ID NO 7561; 650pp + Sequence Listing; English.

XX The present invention provides a number of single exon nucleic acid

XX probes which are derived from genomic sequences expressed in the human

XX brain. They can be used to measure gene expression in brain cell samples,

XX which may enable the diagnosis and improved treatment of nervous system

XX diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,

XX epilepsy and cancers. The present sequence is one of the probes of the

XX invention

XX SQ Sequence 484 BP; 166 A; 95 C; 138 G; 85 T; 0 U; 0 Other;

XX Query Match 80.8%; Score 20.2; DB 4; Length 484;

XX Best Local Similarity 88.0%; Pred. No. 42;

XX Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAAAAAATCAGCCGCGCAGGCATGG 25

DB 55 AAAAAACAAGCAAGCCGCGCAGGCATGG 79

RESULT 38

ABS33124

ID ABS33124 standard; DNA; 484 BP.

XX ABS33124;

XX 25-FEB-2003 (first entry)

XX Human liver single exon probe, SEQ ID NO 8114.

XX Human; single exon nucleic acid probe; liver; cirrhosis;

XX hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;

XX coronary heart disease; ss.

XX Homo sapiens.

OS WO200157275-A2.

PN 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US000667.

XX 04-FEB-2000; 2000US-0180312P.

XX 26-MAY-2000; 2000US-0207456P.

PR 30-JUN-2000; 2000US-00608408.

PR 03-AUG-2000; 2000US-00632366.

PR 21-SEP-2000; 2000US-0234687P.

PR 27-SEP-2000; 2000US-0236359P.

PR 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

PA Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-483446/52.

XX Single exon nucleic acid probes for analyzing gene expression in human

XX brains.

XX Example 4; SEQ ID NO 7561; 650pp + Sequence Listing; English.

XX The present invention provides a number of single exon nucleic acid

XX probes which are derived from genomic sequences expressed in the human

XX brain. They can be used to measure gene expression in brain cell samples,

XX which may enable the diagnosis and improved treatment of nervous system

XX diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,

XX epilepsy and cancers. The present sequence is one of the probes of the

XX invention

XX SQ Sequence 484 BP; 166 A; 95 C; 138 G; 85 T; 0 U; 0 Other;

XX Query Match 80.8%; Score 20.2; DB 4; Length 484;

XX Best Local Similarity 88.0%; Pred. No. 42;

XX Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAAAAAATCAGCCGCGCAGGCATGG 25

DB 55 AAAAAACAAGCAAGCCGCGCAGGCATGG 79

RESULT 38

ABS33124

ID ABS33124 standard; DNA; 484 BP.

XX ABS33124;

XX 25-FEB-2003 (first entry)

XX Human liver single exon probe, SEQ ID NO 8114.

XX Human; single exon nucleic acid probe; liver; cirrhosis;

XX hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;

XX coronary heart disease; ss.

PR 26-MAY-2000; 2000US-0207456P.

PR 30-JUN-2000; 2000US-00608408.

PR 03-AUG-2000; 2000US-00632366.

PR 21-SEP-2000; 2000US-0234687P.

PR 27-SEP-2000; 2000US-0236359P.

PR 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488898/53.

XX Human genome-derived single exon nucleic acid probes useful for analyzing

XX gene expression in human adult liver.

XX Claim 1; SEQ ID NO 8114; 658pp; English.

XX The invention relates to a single exon nucleic acid probe (SENP) (I) for

XX measuring human gene expression in a sample derived from human adult

XX liver, comprising one of 13109 defined nucleotide sequences given in the

XX specification (or complements/ fragments). The probe hybridises at high

XX stringency to a nucleic acid molecule expressed in the human adult liver.

XX (I) may be used for predicting, measuring and displaying gene expression

XX in samples derived from human adult liver. The genes identified may be

XX involved in genetic liver diseases such as cirrhosis,

XX hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is

XX associated with coronary heart disease. ABS25011-ABS51005 represent human

XX liver single exon nucleic acid probes of the invention. Note: The

XX sequence information for this patent does not appear in the printed

XX specification but was obtained in electronic format directly from WIPO at

XX ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 484 BP; 166 A; 95 C; 138 G; 85 T; 0 U; 0 Other;

XX Query Match 80.8%; Score 20.2; DB 4; Length 484;

XX Best Local Similarity 88.0%; Pred. No. 42;

XX Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAAAAAATCAGCCGCGCAGGCATGG 25

DB 55 AAAAAACAAGCAAGCCGCGCAGGCATGG 79

RESULT 39

ABS08206

ID ABS08206 standard; DNA; 484 BP.

XX ABS08206;

XX 19-AUG-2002 (first entry)

XX Human genome-derived single exon probe from lung SEQ ID No 8197.

XX Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;

XX chronic obstructive pulmonary disease; interstitial lung disease;

XX familial idiopathic pulmonary fibrosis; neurofibromatosis;

XX tuberosus sclerosis; Gaucher's disease; Niemann-Pick disease;

XX Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;

XX pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;

XX primary ciliary dyskinesia; fibrocystic pulmonary dysplasia;

XX hyaline membrane disease.

XX Homo sapiens.

XX WO200186003-A2.

XX 15-NOV-2001.

XX 30-JAN-2001; 2001WO-US000665.

XX 04-FEB-2000; 2000US-0180312P.

PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-02346878.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2002-114183/15.
XX Spatially-addressable set of single exon nucleic acid probes, used to
PT measure gene expression in human lung samples.
XX Claim 1; SEQ ID NO 8197; 634pp; English.
XX The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human lung comprising single exon nucleic acid probes having one of
CC 12614 nucleic acid sequences mentioned in the specification, or their
CC complements or the 12387 open reading frames derived from the 12614
CC probes. Also included are a microarray comprising the novel set of probes
CC; the novel set of probes which hybridise at high stringency to a nucleic
CC acid expressed in the human lung; measuring gene expression in a sample
CC derived from human lung, comprising (a) contacting the array with a
CC collection of detectably labeled nucleic acids derived from human lung
CC mRNA, and (b) measuring the label detectably bound to each probe of the
CC array; identifying exons in a eukaryotic genome, comprising (a)
CC algorithmically predicting at least one exon from genomic sequences of
CC the eukaryote; and (b) detecting specific hybridisation of detectably
CC labeled nucleic acids from eukaryotic lung mRNA, to a single exon probe,
CC having a fragment identical to the predicted exon, the probe is included
CC in the above mentioned microarray; assigning exons to a single gene,
CC comprising (a) identifying exons from genomic sequence by the method
CC above and (b) measuring the expression of each of the exons in several
CC tissues and/or cell types using hybridisation to a single exon
CC microarrays having a probe with the exon, where a common pattern of
CC expression of the exons in the tissues and/or cell types indicates that
CC the exons should be assigned to a single gene; a peptide comprising one
CC of 12011 sequences, mentioned in the specification, or encoded by the
CC probes/open reading frames (ORF). The probes are used for gene expression
CC analysis, and for identifying exons in a gene, particularly using human
CC lung derived mRNA and for the study of lung diseases such as asthma, lung
CC cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
CC disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
CC tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
CC Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
CC histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis,
CC Karagenar syndrome, fibrocystic pulmonary dysplasia, primary ciliary
CC dyskinesia, pulmonary hypertension and hyaline membrane disease. The
CC present sequence is a single exon probe of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences
SQ Sequence 484 BP; 166 A; 95 C; 138 G; 85 T; 0 U; 0 Other;
Query Match 80.8%; Score 20.2; DB 6; Length 484;
Best Local Similarity 88.0%; Pred. No. 42;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 AAAAAAAAAATCAGCGCCAGGCATGG 25
Db 55 AAAAAACAAGCAGCGCCAGGCATGG 79
RESULT 40
ID ABN64674
XX ABN64674 standard; cDNA; 526 BP.
XX ABN64674;

XX 28-JUN-2002 (first entry)
XX Human cancer related polynucleotide SEQ ID NO 4641.
DE Human cancer related polynucleotide SEQ ID NO 4641.
XX Human; cytostatic; gene expression; gene mapping; tissue profiling;
KW gene therapy; cancer; tumour; gene; ss.
XX Homo sapiens.
XX WO200214500-A2.
XX 21-FEB-2002.
XX 16-AUG-2001; 2001WO-US025840.
XX 16-AUG-2000; 2000US-0226326P.
XX (CHIR) CHIRON CORP.
PA (HYSE-) HYSEQ INC.
XX Escobedo J, Garcia PD, Sudduth-Klinger J, Reinhard C, Randazzo P;
PI Lamson G, Scott EM, Zhang G, Kassam A, Pot D, Labat I;
XX WPI; 2002-241905/29.
XX New nucleic acid for producing a polypeptide, detecting differentially
PT expressed genes correlated with a cancerous state of a mammalian cell,
PT and inhibiting tumor growth.
XX Claim 1; SEQ ID NO 4641; 883pp + Sequence Listing; English.
XX The invention relates to an isolated polynucleotide (ABN27253-ABN33262)
CC with cytostatic activity. The polynucleotide is used to produce a
CC polypeptide, to detect differentially expressed genes correlated with a
CC cancerous state of a mammalian cell and to inhibit tumour growth. The
CC polynucleotide is used as a probe in mapping and tissue profiling. The
CC encoded polypeptide and antibodies to the polypeptide can also be used
CC for therapeutic and diagnostic purposes. The polynucleotide is useful for
CC gene therapy. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 526 BP; 172 A; 103 C; 120 G; 131 T; 0 U; 0 Other;
Query Match 80.8%; Score 20.2; DB 6; Length 526;
Best Local Similarity 88.0%; Pred. No. 43;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 AAAAAAAAAATCAGCGCCAGGCATGG 25
Db 338 AAAAAAGATTTCAGGCGCCAGGCATGG 362
RESULT 41
ID AAA35120
XX AAA35120 standard; DNA; 666 BP.
XX AAA35120;
XX 28-JUL-2000 (first entry)
XX Human adenosine receptor related polynucleotide SEQ ID NO:2809.
XX Human; adenosine receptor; low adenosine antisense oligonucleotide;
KW phosphorothioate; impaired respiration; inflammation; allergy;
KW allergic disease; bronchoconstriction; inhibitor; antiinflammatory;
KW antiallergic; antiasthmatic; cytostatic; analgesic; impaired airway;
KW lung disease; ischaemic condition; pulmonary vasoconstriction; asthma;
KW respiratory distress syndrome; pain; cystic fibrosis; emphysema;
KW pulmonary hypertension; chronic obstructive pulmonary disease; COPD;
KW cancer; leukaemia; lymphoma; carcinoma; metastasis; ss.
XX

OS Homo sapiens.
PN WO200009525-A2.
XX
XX
PD 24-FEB-2000.
XX
XX 03-AUG-1999; 99WO-US017712.
XX
XX 03-AUG-1998; 98US-0095212P.
XX
XX (UYEC-) UNIV EAST CAROLINA.
PA
XX
XX Nyce JW;
XX
XX WPI; 2000-205971/18.
XX
XX New antisense oligonucleotides useful for treating e.g. pulmonary
PT vasoconstriction, inflammation, allergies, asthma, hypertension,
PT bronchitis, emphysema, respiratory distress syndrome, ischemia or
PT cancers.
XX
XX Disclosure; Page 1075-1076; 1343pp; English.
XX
XX The present invention describes a new composition comprising an antisense
CC oligonucleotide (ON) with low adenine (up to 15%), which targets
CC nucleic acids involved in bronchoconstriction, allergies, and/or
CC inflammation. The ON can have antiinflammatory, antiallergic,
CC antiasthmatic, cytostatic and analgesic activities. The compositions are
CC useful for the treatment of diseases associated with inflammation,
CC impaired airways, including lung disease and diseases whose secondary
CC effects afflict the lungs of a subject. They can be used for treating
CC e.g. ischaemic conditions, pulmonary vasoconstriction, allergies, asthma,
CC impaired respiration, respiratory distress syndrome, pain, cystic
CC fibrosis, pulmonary hypertension, emphysema, chronic obstructive
CC pulmonary disease (COPD), and cancers such as leukaemias, lymphomas,
CC carcinomas, and cancers which may metastasise to the lungs, including
CC breast and prostate cancer. The reduction of the adenine content of the
CC ONs reduces side effects. The A-containing ONs break down with the
CC release of deoxyadenosine which activates adenosine receptors causing
CC bronchoconstriction and inflammation. AAA32313 to AAA35312 represent the
CC nucleotide sequences given in the sequence listing from the present
CC invention, which correspond to SEQ ID NO:1 to 2815, and then the last 185
CC sequences are also called SEQ ID NO:1 to 185, but the sequences differ
CC from the previously named sequences. SEQ ID NO:11 to 1690 (AAA32323 to
CC AAA33992) are specifically claimed ONs from the present invention. N.B.
CC Sequences given in the disclosure of the present invention do not match
CC up with their corresponding SEQ ID NO: sequences given in the sequence
CC listing
XX
XX Sequence 666 BP; 206 A; 165 C; 174 G; 121 T; 0 U; 0 Other;
SQ
Query Match 80.8%; Score 20.2; DB 3; Length 666;
Best Local Similarity 88.0%; Pred. No. 44;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 1 AAAAAAAAAATCAGCCAGCATGG 25
Db 99 AAAAAAAAAAAGCCAGCATGG 123
RESULT 42
AAF21242
ID AAF21242 standard; DNA; 666 BP.
XX
AC AAF21242;
XX
XX 14-MAR-2001 (first entry)
XX
XX Human low adenine antisense oligonucleotide related sequence #2809.
DE
XX Low adenine antisense oligonucleotide; phosphorothioate; allergy;
XX human; airway disorder; bronchoconstriction; lung inflammation;
XX surfactant depletion; respiratory; bronchodilator; antiinflammatory;

KW immunosuppressive; antiaethmatic; analgesic; hypotensive; cytostatic;
KW respiratory obstruction; pulmonary vasoconstriction; impeded respiration;
KW surfactant hypoproduction; pulmonary vasoconstriction; asthma; RDS;
KW respiratory distress syndrome; pain; cystic fibrosis; allergic rhinitis;
KW pulmonary hypertension; emphysema; pulmonary transplantation rejection;
KW chronic obstructive pulmonary disease; pulmonary infection; bronchitis;
KW cancer; ss.
XX
XX Homo sapiens.
XX
XX WO2000062736-A2.
PN
XX
XX 26-OCT-2000.
XX
XX 24-MAR-2000; 2000WO-US008020.
XX
XX 06-APR-1999; 99US-0127958P.
XX
XX (UYEC-) UNIV EAST CAROLINA.
PA
XX (NYCE/) NYCE J W.
XX
XX Nyce JW;
XX
XX WPI; 2000-679539/66.
XX
XX Low adenine (A) content antisense oligonucleotides which do not trigger
PT adenosine receptors during metabolism, useful e.g. for treating cancers
PT and respiratory obstructions.
XX
XX Disclosure; Page 1155; 1592pp; English.
XX
XX The present invention describes low adenine (A) content antisense
CC oligonucleotides and compositions (I) comprising them. In the antisense
CC oligonucleotides the A is replaced by a 'Universal' or alternative base.
CC (I) can have respiratory, bronchodilator, antiinflammatory, analgesic,
CC immunosuppressive, antiaethmatic, hypotensive and cytostatic activities.
CC The antisense oligonucleotides and (I) can be used to down-regulate the
CC expression and/or activity of target polypeptides associated with
CC lung/respiratory disorders and malignancies, such as stimulating and
CC activating peptide factors and transmitters, transcription factors,
CC immunoglobulins and antibodies, antibody receptors, cytokines and
CC chemokines, endogenously produced specific and non-specific enzymes,
CC binding proteins, adhesion molecules and their receptors, cytokine and
CC chemokine receptors, adenosine receptors, bradykinin receptors, central
CC nervous system (CNS) and peripheral nervous and non-nervous system
CC receptors, CNS and peripheral nervous and non-nervous system peptide
CC transmitters, defensins, growth factors, vasoactive peptides and
CC receptors, binding proteins and malignancy associated proteins. The
CC antisense oligonucleotides may be used in this way to treat disorders
CC including respiratory obstruction (especially pulmonary obstruction
CC and/or bronchoconstriction) and/or lung inflammation, allergy(ies) and/or
CC surfactant hypoproduction which are associated with a disease or
CC condition selected from pulmonary vasoconstriction, inflammation,
CC allergies, asthma, impeded respiration, respiratory distress syndrome
CC (RDS), pain, cystic fibrosis (CF), allergic rhinitis (AR), pulmonary
CC hypertension, emphysema, chronic obstructive pulmonary disease (COPD),
CC pulmonary transplantation rejection, pulmonary infections, bronchitis,
CC and/or cancer. AAF18434 to AAF21543 represent human polynucleotide
CC fragments and antisense oligonucleotides used in the exemplification of
CC the present invention
XX
XX Sequence 666 BP; 206 A; 165 C; 174 G; 121 T; 0 U; 0 Other;
SQ
Query Match 80.8%; Score 20.2; DB 3; Length 666;
Best Local Similarity 88.0%; Pred. No. 44;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 1 AAAAAAAAAATCAGCCAGCATGG 25
Db 99 AAAAAAAAAAAGCCAGCATGG 123
RESULT 43

```
ABZ96936
ID ABZ96936 standard; DNA; 666 BP.
XX
AC ABZ96936;
XX
XX 17-OCT-2003 (first entry)
XX
XX Human nucleic acid sequence.
XX
XX Human; antisense; lung dysfunction; nasal airway dysfunction;
KW antinflammatory steroid; ubiqunone; antiinflammatory; antiallergic;
KW antiasthmatic; hypotensive; immunosuppressive; cytostatic; gene therapy;
KW antisense gene therapy; respiratory; lung; adenosine sensitivity;
KW adenosine receptor; bronchodilation; bronchoconstriction; lung allergy;
KW lung inflammation; respiratory disease; ds.
XX
XX Homo sapiens.
OS
XX WO200285308-A2.
PN
XX 31-OCT-2002.
PD
XX 23-APR-2002; 2002WO-US013135.
PF
XX 24-APR-2001; 2001US-0286137P.
PR
XX (EPITG-) EPIGENESIS PHARM INC.
XX
XX Nyce JW, Li Y, Sandrasagra A, Katz E, Pabalan J, Aguilar D;
PI Miller S, Tang L, Shahabuddin S;
XX WPI; 2003-229219/22.
XX
XX Pharmaceutical composition for treating ailments associated with impaired
PT respiration, has oligo(s) antisense to specific gene(s) or its
PT corresponding RNAs, and glucocorticoid or non-glucocorticoid steroid or
PT ubiqunone.
XX
XX Disclosure; SEQ ID NO 12178; 872pp; English.
XX
XX The invention relates to a novel pharmaceutical composition, which has a
CC first active agent comprising an oligonucleotide antisense to the
CC initiation codon, coding region, 5' or 3' end genomic flanking regions,
CC 5' and 3' intron-exon junctions, or regions within 2-10 nucleotides of
CC junctions of genes encoding a polypeptide associated with lung and/or
CC nasal airway dysfunction and a second active agent comprising an
CC antinflammatory steroid and ubiqunone. A composition of the invention
CC has antinflammatory, antiallergic, antiasthmatic, hypotensive,
CC immunosuppressive, and cytostatic activity. The composition may have a
CC use in antisense gene therapy. The composition is useful for treating or
CC preventing a respiratory, lung or malignant disease or condition, also
CC for enhancing the prophylactic or therapeutic respiratory effect of an
CC antinflammatory steroid in a subject, for reducing or depleting levels
CC of, or reducing sensitivity to adenosine, reducing levels of adenosine
CC receptor, producing bronchodilation, increasing levels of ubiqunone or
CC lung surfactant in a subject's tissue, or treating bronchoconstriction,
CC lung inflammation, lung allergies, or a respiratory disease or condition.
CC Note: The sequence data for this patent is not represented in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 666 BP; 206 A; 165 C; 174 G; 121 T; 0 U; 0 Other;
SQ
Query Match 80.8%; Score 20.2; DB 10; Length 666;
Best Local Similarity 88.0%; Pred. No. 44;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 AAAAAAAAAATCAGCGCCAGCGATGG 25
| | | | | | | | | | | | | | | | | | | | |
Db 99 AAAAAAAAAAAAAAGGCGCAGCGATGG 123
| | | | | | | | | | | | | | | | | | | | |
RESULT 44
```

```
SQ Sequence 666 BP; 206 A; 165 C; 174 G; 121 T; 0 U; 0 Other;
Query Match      80.8%; Score 20.2; DB 11; Length 666;
Best Local Similarity 88.0%; Pred. No. 44;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAAAAAATCAGCGCAGCATGG 25
Db 99 AAAAAAAAAAAGCCAGGCATGG 123

RESULT 45
AAS28505/c
ID AAS28505 standard; DNA; 1131 BP.
XX
AC AAS28505;
XX
DT 07-NOV-2001 (first entry)
XX
DE Genomic sequence #345 encoding for novel human respiratory antigen.
XX
KW Human; respiratory antigen; respiratory disorder; throat disorder;
KW lung disorder; nose disorder; lung cancer; gene therapy; cytostatic;
KW anti allergic; anti asthmatic; anti inflammatory; oilyfactory;
KW respiratory active; ds.
XX
OS Homo sapiens.
XX
PN WO20015448-A1.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US0001333.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 18-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0232968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239355P.
PR 13-OCT-2000; 2000US-0239377P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
```



```
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 23-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 17-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 06-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-457723/49.
XX
XX Isolated polypeptide for treating, preventing and/ or prognosing
XX respiratory disorders related to the lung including lung cancers and also
XX for testing and detection e.g. diagnosis.
XX
XX Claim 1; SEQ ID NO 260; 507pp; English.
XX
XX Sequences AAS29931-AAS30164 represent genomic DNA molecules, which encode
XX the lung antigen polypeptides of the invention. Lung antigen polypeptides
XX and their associated polynucleotides are useful in the diagnosis,
XX treatment and prevention of various types of disorders in e.g. humans,
XX mice, rabbits, goats, horses, cats, dogs, chickens or sheep. A
XX pathological condition can be determined by detecting the presence or
XX absence of a mutation in a lung antigen polynucleotide. The treatable
XX disorders include autoimmune diseases such as neoplasms of the breast or liver,
XX hyperproliferative disorders such as cardiac arrest, cerebrovascular
XX cardiovascular disorders such as cerebral ischaemia, nervous system disorders such as
XX Alzheimer's disease, infections caused by bacteria, viruses and fungi,
XX ocular disorders such as corneal infection, gastrointestinal disorders such as
XX premature labour and infertility, gastrointestinal disorders such as
XX Crohn's disease, renal disorders such as asthma and pleurisy. The polypeptides can
XX also be used to aid wound healing, to prevent skin aging due to sunburn,
XX to maintain organs before transplantation, to regenerate tissues and in
XX chemotaxis. The polypeptides can also be used as a food additive or
XX preservative to increase or decrease storage capabilities. Note: The
XX sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences
XX
XX Query Match 80.8%; Score 20.2; DB 5; Length 1131;
XX Best Local Similarity 88.0%; Pred. No. 48;
XX Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Oy 1 AAAAAAAAAACAGGCCAGCGCATGG 25
Db 500 AAAAAAAAAATTAAGGCCAGCGCATGG 476
RESULT 47
ADB33333/c
ID ADB33333 standard; DNA; 1131 BP.
XX
XX ADB33333;
XX
XX 04-DEC-2003 (first entry)
XX
XX Human novel lung related polypeptide DNA SEQ ID NO 260.
XX
XX gene therapy; lung antigen; neoplasia; acute myelogenous leukaemia;
XX adenocarcinoma; respiratory disorder; chronic rhinitis; sinusitis;
XX immunodeficiency; X-linked agammaglobulinaemia;
XX X-linked infantile agammaglobulinaemia; inflammatory disorder;
XX adrenailitis; alveolitis; immune complex disease; serum sickness;
XX polyarteritis nodosa; bleeding disorder; thrombocytopenia;
XX Von Willebrand's disease; acquired platelet dysfunction; kidney failure;
```

multiple myeloma; macrophage related disorder; Gaucher's disease;
Neimann-Pick disease, tumour; colon cancer; pancreatic cancer;
renal disorder; nephritis; bone disorder; Albers-Schönberg disease;
bowleg; muscle disorder; Becker's muscular dystrophy;
Duchenne's muscular dystrophy; nervous disorder; ischaemic lesion;
traumatic lesion; endocrine disorder; Cushing's syndrome;
corticosteroid deficiency; gastrointestinal disorder; dysphagia;
gastric reflux; human; ds.

Homo sapiens.

US2003054368-A1.

20-MAR-2003.

22-FEB-2002; 2002US-00079854.

31-JAN-2000; 2000US-0179065P.

04-FEB-2000; 2000US-0180628P.

24-FEB-2000; 2000US-0184664P.

02-MAR-2000; 2000US-0186350P.

16-MAR-2000; 2000US-0189874P.

17-MAR-2000; 2000US-0190076P.

18-APR-2000; 2000US-0198123P.

19-MAY-2000; 2000US-0205515P.

07-JUN-2000; 2000US-0209467P.

28-JUN-2000; 2000US-0214886P.

30-JUN-2000; 2000US-0215135P.

07-JUL-2000; 2000US-0216647P.

07-JUL-2000; 2000US-0216680P.

11-JUL-2000; 2000US-0217487P.

11-JUL-2000; 2000US-0217496P.

14-JUL-2000; 2000US-0218290P.

26-JUL-2000; 2000US-0220963P.

26-JUL-2000; 2000US-0220964P.

14-AUG-2000; 2000US-0224518P.

14-AUG-2000; 2000US-0224519P.

14-AUG-2000; 2000US-0225213P.

14-AUG-2000; 2000US-0225214P.

14-AUG-2000; 2000US-0225266P.

14-AUG-2000; 2000US-0225267P.

14-AUG-2000; 2000US-0225268P.

14-AUG-2000; 2000US-0225270P.

14-AUG-2000; 2000US-0225447P.

14-AUG-2000; 2000US-0225757P.

14-AUG-2000; 2000US-0225758P.

14-AUG-2000; 2000US-0225759P.

18-AUG-2000; 2000US-0226279P.

22-AUG-2000; 2000US-0226681P.

22-AUG-2000; 2000US-0226686P.

22-AUG-2000; 2000US-0227182P.

23-AUG-2000; 2000US-0227009P.

30-AUG-2000; 2000US-0228924P.

01-SEP-2000; 2000US-0229287P.

01-SEP-2000; 2000US-0229343P.

01-SEP-2000; 2000US-0229344P.

01-SEP-2000; 2000US-0229345P.

05-SEP-2000; 2000US-0229509P.

05-SEP-2000; 2000US-0229513P.

06-SEP-2000; 2000US-0230437P.

06-SEP-2000; 2000US-0230438P.

08-SEP-2000; 2000US-0231242P.

08-SEP-2000; 2000US-0231243P.

08-SEP-2000; 2000US-0231244P.

08-SEP-2000; 2000US-0231413P.

08-SEP-2000; 2000US-0231414P.

08-SEP-2000; 2000US-0232080P.

08-SEP-2000; 2000US-0232081P.

12-SEP-2000; 2000US-0231968P.

14-SEP-2000; 2000US-0232397P.

14-SEP-2000; 2000US-0232398P.

14-SEP-2000; 2000US-0232399P.

14-SEP-2000; 2000US-0232400P.

14-SEP-2000; 2000US-0232401P.

14-SEP-2000; 2000US-0233063P.

14-SEP-2000; 2000US-0233064P.

14-SEP-2000; 2000US-0233065P.

21-SEP-2000; 2000US-0234223P.

21-SEP-2000; 2000US-0234274P.

25-SEP-2000; 2000US-0234997P.

25-SEP-2000; 2000US-0234998P.

26-SEP-2000; 2000US-0235484P.

27-SEP-2000; 2000US-0235834P.

27-SEP-2000; 2000US-0235836P.

29-SEP-2000; 2000US-0236327P.

29-SEP-2000; 2000US-0236372P.

29-SEP-2000; 2000US-0236368P.

29-SEP-2000; 2000US-0236369P.

29-SEP-2000; 2000US-0236370P.

02-OCT-2000; 2000US-0236802P.

02-OCT-2000; 2000US-0237037P.

02-OCT-2000; 2000US-0237038P.

02-OCT-2000; 2000US-0237039P.

02-OCT-2000; 2000US-0237040P.

13-OCT-2000; 2000US-0239933P.

13-OCT-2000; 2000US-0239937P.

20-OCT-2000; 2000US-0240960P.

20-OCT-2000; 2000US-0241221P.

20-OCT-2000; 2000US-0241785P.

20-OCT-2000; 2000US-0241786P.

20-OCT-2000; 2000US-0241787P.

20-OCT-2000; 2000US-0241808P.

20-OCT-2000; 2000US-0241809P.

20-OCT-2000; 2000US-0241826P.

01-NOV-2000; 2000US-0244617P.

08-NOV-2000; 2000US-0246474P.

08-NOV-2000; 2000US-0246475P.

08-NOV-2000; 2000US-0246476P.

08-NOV-2000; 2000US-0246477P.

08-NOV-2000; 2000US-0246478P.

08-NOV-2000; 2000US-0246521P.

08-NOV-2000; 2000US-0246524P.

08-NOV-2000; 2000US-0246525P.

08-NOV-2000; 2000US-0246526P.

08-NOV-2000; 2000US-0246527P.

08-NOV-2000; 2000US-0246528P.

08-NOV-2000; 2000US-0246532P.

08-NOV-2000; 2000US-0246603P.

08-NOV-2000; 2000US-0246610P.

08-NOV-2000; 2000US-0246611P.

08-NOV-2000; 2000US-0246613P.

17-NOV-2000; 2000US-0249207P.

17-NOV-2000; 2000US-0249208P.

17-NOV-2000; 2000US-0249209P.

17-NOV-2000; 2000US-0249210P.

17-NOV-2000; 2000US-0249211P.

17-NOV-2000; 2000US-0249212P.

17-NOV-2000; 2000US-0249213P.

17-NOV-2000; 2000US-0249214P.

17-NOV-2000; 2000US-0249215P.

17-NOV-2000; 2000US-0249216P.

17-NOV-2000; 2000US-0249217P.

17-NOV-2000; 2000US-0249218P.

17-NOV-2000; 2000US-0249244P.

17-NOV-2000; 2000US-0249245P.

17-NOV-2000; 2000US-0249264P.

17-NOV-2000; 2000US-0249265P.

17-NOV-2000; 2000US-0249297P.

17-NOV-2000; 2000US-0249299P.

01-DEC-2000; 2000US-0249300P.

01-DEC-2000; 2000US-0250160P.

01-DEC-2000; 2000US-0250391P.

05-DEC-2000; 2000US-0251030P.

05-DEC-2000; 2000US-0251988P.

05-DEC-2000; 2000US-0256719P.

06-DEC-2000; 2000US-0251479P.

PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251866P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
PR 17-JAN-2001; 2001US-00764878.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM, Barash SC;
XX
DR WPI; 2003-695900/66.
XX
XX Novel isolated lung antigen polypeptides useful for treating, preventing,
PT diagnosing acute myelogenous leukemias, adenocarcinoma, thrombocytopenia,
PT Von Willebrand's disease.
XX
XX Disclosure; SEQ ID NO 260; 178pp; English.
XX
XX The invention relates to an isolated lung antigen polypeptide sequence or
CC encoded sequence in a cDNA clone. The polypeptide and its polynucleotide
CC are useful for treating, preventing, diagnosing and/or prognosing
CC diseases and/or disorders such as pathological cell proliferative
CC neoplasias e.g. acute myelogenous leukaemias, adenocarcinoma; respiratory
CC disorders such as chronic rhinitis, sinusitis; immunodeficiencies such as
CC X-linked agammaglobulinaemia, X-linked infantile agammaglobulinaemia;
CC inflammatory disorders such as adrenailitis, alveolitis; immune complex
CC diseases such as serum sickness, polyarteritis nodosa; bleeding disorders
CC such as thrombocytopenia, Von Willebrand's disease; acquired platelet
CC dysfunction such as kidney failure, multiple myeloma; disorders
CC associated with macrophage numbers and/or macrophage function such as
CC Gaucher's disease, Niemann-Pick disease; tumours such as colon cancer,
CC pancreatic cancer; renal disorders such as kidney failure, nephritis;
CC bone disorders such as Albers-Schonberg disease, bowlegs; muscle
CC disorders such as Becker's muscular dystrophy, Duchenne's muscular
CC dystrophy; nervous disorders such as ischaemic lesions, traumatic lesions
CC ; endocrine disorders such as Cushing's syndrome, corticosteroid

Query Match 80.8%; Score 20.2; DB 10; Length 1131;
Best Local Similarity 88.0%; Pred. No. 48;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AAAAAAAAAATCAGGCCAGGCATGG 25
||||||| | |||||
Db 500 AAAAAAAAAATTAAGGCCAGGCATGG 476

RESULT 48
ADG41701/c
ID ADG41701 standard; DNA; 1131 BP.
XX
AC ADG41701;
XX
DT 26-FEB-2004 (first entry)
XX
DE Human respiratory system associated genomic DNA seq id 939.
XX
XX antiinflammatory; antiallergic; antiasthmatic; cytostatic; gene therapy;
KW respiratory system antigen;
KW human respiratory system associated polynucleotide;
KW respiratory system disorder; throat disorder; vocal cord paralysis;
KW tonsillitis; laryngitis; lung disorder; pneumonia; allergic disorder;
KW asthma; eosinophilic pneumonia; pleurisy; cystic fibrosis; emphysema;
KW histiocytosis; sarcoidosis; nose disorder; rhinitis; sinusitis; neoplasm;
KW cancer; respiratory tissue cancer; throat cancer; lung cancer;
KW cancer of the nose; gene therapy; chromosome identification; forensic;
KW human respiratory system associated protein; ds; human.
XX
OS Homo sapiens.
XX
XX US2003215893-A1.
FN

XX 20-NOV-2003.
PD
XX 07-AUG-2002; 2002US-00212872.
PF
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 28-JUN-2000; 2000US-0209467P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 23-AUG-2000; 2000US-0226968P.
PR 23-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229145P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 12-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 21-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 28-SEP-2000; 2000US-0235935P.

PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0241886P.
PR 30-JUN-2000; 2000US-0215113P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 03-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232337P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.

PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 01-DEC-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 05-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
PR 17-JAN-2001; 2001US-00764860.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Ruben SM, Barash SC;

WPI; 2003-765403/72.

New human respiratory system-related polypeptide and genes, useful for treating, preventing or diagnosing e.g. pneumonia, lung cancer, cystic fibrosis, asthma, sarcoidosis, rhinitis, leukemia, inflammations or sinusitis.

Disclosure; SEQ ID NO 939; 202pp; English.

This invention is related to a novel isolated polypeptide, which comprises a human respiratory system-related polypeptide, and the DNA

CC sequence which encodes it. The invention may be useful for the
CC development of compounds with an antiasthmatic, antibacterial
CC antiinflammatory, cytostatic, antianaemic or antiallergic activity. In
CC addition, the sequences disclosed may be useful for gene therapy. The
CC polypeptide or polynucleotide is useful for treating, preventing or
CC ameliorating a medical condition, for example pneumonia, lung cancer,
CC cystic fibrosis, asthma, sarcoidosis, rhinitis, anaemia, leukaemia,
CC inflammations, sinusitis, chronic obstructive pulmonary disease or
CC infectious diseases. The polypeptide or polynucleotide is also useful for
CC diagnosing any of these diseases or a susceptibility to the disease. The
CC present sequence is that of a human DNA sequence which is related to a
CC human respiratory system associated gene of the invention.
XX
XX
SQ Sequence 1131 BP; 311 A; 240 C; 271 G; 309 T; 0 U; 0 Other;

Query Match 80.8%; Score 20.2; DB 11; Length 1131;
Best Local Similarity 88.0%; Pred. No. 48;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAAAAAATCAGGCCAGGCATGG 25
||||||| | | | | | | | | | |
DB 500 AAAAAAAAAATTAAGGCCAGGCATGG 476

RESULT 50
ABQ75320
ID ABQ75320 standard; cDNA; 1264 BP.
XX
AC ABQ75320;
XX
DT 05-NOV-2002 (first entry)
XX
DE Human lung specific nucleic acid sequence SEQ ID NO:59.
XX
KW Human; lung; lung specific nucleic acid; LSNA; lung specific protein;
KW LSP; cytostatic; gene therapy; vaccine; metastasis; lung cancer;
KW squamous cell carcinoma; gene; chromosome 3; ss.
XX
OS Homo sapiens.
XX
PN WO200264788-A2.
XX
PD 22-AUG-2002.
XX
PF 20-NOV-2001; 2001WO-US045080.
XX
PR 20-NOV-2000; 2000US-0252054P.
XX
PA (DIAD-) DIADEXUS INC.
XX
PI Macina RA, Recipon H, Chen S, Sun Y, Liu C;
XX
XX WPI; 2002-657601/70.

XX
XX
PT New lung specific nucleic acid useful in gene therapy or as vaccines for
PT treating lung cancer (e.g. squamous cell carcinoma) or non-carcinoma lung
PT diseases, as well as for diagnosing, monitoring or staging these
PT diseases.

XX
XX
PS Claim 1; Page 181-182; 282pp; English.

XX
XX
XX The present invention describes an isolated lung specific nucleic acid
XX (LSNA) comprising a sequence that: (a) encodes any of the 93 amino acid
XX sequences comprising 17 - 733 amino acids, given in ABP52873 to ABP52965;
XX (b) comprises any of 115 sequences comprising 148 - 3193 base pairs (bp),
XX given in ABQ75362 to ABQ75376; (c) selectively hybridises to (a) or (b);
XX or (d) has 60 % sequence identity to (a) or (b). LSNA and lung specific
XX protein (LSP) sequences have cytostatic activity and can be used in gene
XX therapy and vaccines. LSNA and LSPs are useful for diagnosing and
XX monitoring the presence and metastases of lung cancer in a patient. An
XX antibody that specifically binds to an LSP can be used for determining
XX the presence of an LSP in a sample, as well as for treating a patient
XX with lung cancer, particularly by inducing an immune response against the

CC lung cancer cell expressing the LSNA or LSPs. In particular, these LSNA
CC and LSPs are useful for identifying, diagnosing, monitoring, staging,
CC imaging and treating lung cancer (e.g. squamous cell carcinoma) and non-
CC cancerous disease states in lung

XX
SQ Sequence 1264 BP; 318 A; 315 C; 282 G; 349 T; 0 U; 0 Other;

Query Match 80.8%; Score 20.2; DB 6; Length 1264;
Best Local Similarity 88.0%; Pred. No. 49;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAAAAAATCAGGCCAGGCATGG 25
||||||| | | | | | | | | | |
DB 485 AAAAAAAAAATTAAGGCCAGGCATGG 509

Search completed: February 3, 2006, 21:56:15
Job time : 212.111 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.
OM nucleic - nucleic search, using sw model
Run on: February 3, 2006, 14:38:25 ; Search time 2951.67 Seconds
(without alignments)
396.277 Million cell updates/sec
Title: US-10-719-900-1
Perfect score: 25
Sequence: 1 aaaaaaaatcacggccagcatgg 25
Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 150 summaries

Database : EST:
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	25	100.0	246	1	BB056745
2	25	100.0	329	1	AW413970
3	25	100.0	410	1	AW763628
4	25	100.0	469	7	CK343886
5	25	100.0	534	2	BB693815
6	25	100.0	649	6	CB248873
7	23.4	93.6	354	8	T69797
8	23.4	93.6	389	1	AA507708
9	22.4	89.6	413	9	AZ008849
10	21.8	87.2	254	1	AW554415
11	21.8	87.2	317	2	BP465089
12	21.8	87.2	322	1	BB105780
13	21.8	87.2	330	7	CB671031
14	21.8	87.2	373	5	BY690105
15	21.8	87.2	373	1	AW492561
16	21.8	87.2	375	2	BB791462
17	21.8	87.2	379	5	BY698909
18	21.8	87.2	386	7	CM698559
19	21.8	87.2	388	5	BY425365
20	21.8	87.2	392	5	BY684912
21	21.8	87.2	394	5	BY691014
22	21.8	87.2	401	5	BY701163

C 23	21.8	87.2	403	5	BY694371
C 24	21.8	87.2	404	5	BY690774
C 25	21.8	87.2	415	5	BY655013
C 26	21.8	87.2	418	8	BG074809
C 27	21.8	87.2	418	8	W12097
C 28	21.8	87.2	421	1	AI124789
C 29	21.8	87.2	423	5	BY697032
C 30	21.8	87.2	425	2	BB781154
C 31	21.8	87.2	425	2	BB789804
C 32	21.8	87.2	426	5	BY383430
C 33	21.8	87.2	431	1	AA744879
C 34	21.8	87.2	431	1	AI642824
C 35	21.8	87.2	434	5	BY693822
C 36	21.8	87.2	436	5	BY384796
C 37	21.8	87.2	447	2	BB782484
C 38	21.8	87.2	470	2	BB819202
C 39	21.8	87.2	474	3	BM205156
C 40	21.8	87.2	491	9	AZ450659
C 41	21.8	87.2	514	3	BM203574
C 42	21.8	87.2	519	1	AA721682
C 43	21.8	87.2	555	2	BB753963
C 44	21.8	87.2	575	3	BP768879
C 45	21.8	87.2	593	7	CK345513
C 46	21.8	87.2	595	1	AW543397
C 47	21.8	87.2	600	3	BP769047
C 48	21.8	87.2	630	10	AG075591
C 49	21.8	87.2	666	1	AW976597
C 50	21.8	87.2	715	7	CK787496
C 51	21.8	87.2	788	6	CB320467
C 52	21.8	87.2	821	8	CK757930
C 53	21.8	87.2	962	5	BA461680
C 54	21.8	87.2	1030	1	AL579068
C 55	21.8	87.2	3341	4	CR749847
C 56	21.8	87.2	3355	4	AK052299
C 57	21.8	87.2	1013	1	AL582650
C 58	20.8	83.2	345	9	AO67376
C 59	20.8	83.2	355	1	AI420666
C 60	20.8	83.2	396	9	AO285592
C 61	20.8	83.2	448	3	BJ582558
C 62	20.8	83.2	525	9	AO373355
C 63	20.8	83.2	675	10	AG165060
C 64	20.8	83.2	709	9	AO487377
C 65	20.8	83.2	719	10	CZ301420
C 66	20.8	83.2	728	9	AO056247
C 67	20.8	83.2	835	9	BZ216878
C 68	20.8	83.2	1068	10	CZ390223
C 69	20.4	81.6	364	2	BF812239
C 70	20.4	81.6	398	2	BF811957
C 71	20.4	81.6	401	2	BF812248
C 72	20.2	80.8	141	9	AZ104693
C 73	20.2	80.8	159	2	BG952210
C 74	20.2	80.8	155	6	CD172275
C 75	20.2	80.8	181	1	AA830764
C 76	20.2	80.8	210	1	AA669747
C 77	20.2	80.8	243	1	AU059999
C 78	20.2	80.8	264	7	CN277450
C 79	20.2	80.8	268	9	AO267741
C 80	20.2	80.8	278	1	AW903860
C 81	20.2	80.8	278	2	BF903277
C 82	20.2	80.8	283	3	BM510815
C 83	20.2	80.8	286	1	AW833004
C 84	20.2	80.8	289	3	BI491315
C 85	20.2	80.8	289	1	AW020255
C 86	20.2	80.8	289	1	AW022442
C 87	20.2	80.8	289	1	AW022611
C 88	20.2	80.8	289	3	BI491894
C 89	20.2	80.8	289	3	BI492396
C 90	20.2	80.8	291	1	AW021440
C 91	20.2	80.8	294	7	CK300314
C 92	20.2	80.8	296	2	BB340383
C 93	20.2	80.8	298	2	BB488258
C 94	20.2	80.8	304	1	BB086006
C 95	20.2	80.8	315	1	BB098444

BY694371	BY694371
BY690774	BY690774
BY655013	BY655013
BG074809	H3139F01-
W12097	ma96f07.r1
AI124789	am56b10.x
BY697032	BY697032
BB781154	BB781154
BB789804	BB789804
BY383430	BY383430
AA744879	nv71h11.s
AI642824	vd62a08.x
BY693822	BY693822
BY384796	BY384796
BB782484	BB782484
BB819202	BB819202
BM205156	CO279F09-
AZ450659	1M0249D01
BM203574	CO256E04-
AA721682	ny72h09.s
BB753963	BB753963
BP768879	BP768879
CK345513	L0247D01-
AW543397	L0240805-
BP769047	BP769047
AG075591	Pan trogl
AW976597	EST388706
CK787496	AGENCOURT
CB320467	AGENCOURT
CK757930	AGENCOURT
BA461680	BA461680
AL579068	AL579068
CR749847	Homo sapi
AK052299	Mus muscu
AL582650	AL582650
AO67376	HS_2237.A
AI420666	tf13c08.x
AO285592	RPC111-89
BJ582558	BJ582558
AO373355	CIT-HSP-2
AG165060	Pan trogl
AO487377	RPC1-11-2
CZ301420	ZMMBF0076
AO056247	CIT-HSP-2
BZ216878	CH230-467
CZ390223	ZMMBF0167
BF812239	QVI-C1017
BF811957	QVI-C1017
BF812248	QVI-C1017
AZ104693	RPCI-23-2
BG952210	CM4-CT062
CD172275	AGENCOURT
AA830764	OC53f07.8
AA669747	adale04.8
AU059999	AU059999
CN277450	170060001
AO267741	RPCI11-69
AW903860	CM4-NN103
BF903277	CM2-MT019
BM510815	ij3e02.x
AW833004	RK3-TT000
BI491315	df07a08.w
AW020255	df07a08.y
AW022442	df38f12.y
AW022611	df41c11.y
BI491894	df16f02.x
BI492396	df23f07.w
AW021440	df23f07.y
CK300314	UI-E-EJ1-
BB340383	BB340383
BB488258	BB488258
BB086006	BB086006
BB098444	BB098444

96	20.2	80.8	319	1	AW903862	AW903862 CM4-NN103
97	20.2	80.8	325	8	T40452	Y401e07.s2
98	20.2	80.8	327	2	BF858276	RC5-FT019
c 99	20.2	80.8	328	2	BF858276	RC5-FT019
c 100	20.2	80.8	335	9	CG061343	MUGQ CH25
c 101	20.2	80.8	336	7	CR773139	DKFZP470B
c 102	20.2	80.8	359	2	BI273543	F4_1A2_Fe
c 103	20.2	80.8	361	5	BY662197	BY662197
c 104	20.2	80.8	375	1	AM168624	xi91e04.x
c 105	20.2	80.8	379	2	BF838333	RC1-HT097
c 106	20.2	80.8	381	7	CR774415	DKFZP459E
c 107	20.2	80.8	381	9	BI1295	CIT-HSP-216
c 108	20.2	80.8	387	8	H96121	Yv18a04.r1
c 109	20.2	80.8	389	9	AQ096814	HS_3039_A
c 110	20.2	80.8	390	5	BY425804	BY425804
c 111	20.2	80.8	393	5	BY700469	BY700469
c 112	20.2	80.8	394	2	BB712180	BB712180
c 113	20.2	80.8	395	2	BG979466	CM4-CN006
c 114	20.2	80.8	395	5	BY700192	BY700192
c 115	20.2	80.8	397	1	AA584686	RC09c07_8
c 116	20.2	80.8	398	9	AQ549431	RPCI-11-4
c 117	20.2	80.8	399	5	BY697747	BY697747
c 118	20.2	80.8	404	8	N66784	Yz56d06.s1
c 119	20.2	80.8	422	8	T90285	Yd42c08.s1
c 120	20.2	80.8	423	1	AA218835	ZQ15a11.s
c 121	20.2	80.8	432	5	BY443017	BY443017
c 122	20.2	80.8	433	1	AI580222	AI580222
c 123	20.2	80.8	434	2	BB783741	BB783741
c 124	20.2	80.8	436	5	BX114796	BX114796
c 125	20.2	80.8	437	1	AW317062	XR26a04.x
c 126	20.2	80.8	440	9	AQ465247	HS_5072_B
c 127	20.2	80.8	443	7	CN421267	1700042Q
c 128	20.2	80.8	443	7	AQ037022	CIT-HSP-2
c 129	20.2	80.8	444	5	BX476002	DKFZP686M
c 130	20.2	80.8	444	9	AQ017348	CIT-HSP-2
c 131	20.2	80.8	456	6	CD236938	FNPAEF06
c 132	20.2	80.8	461	5	BY381322	BY381322
c 133	20.2	80.8	462	1	AA197019	zp09e03.r
c 134	20.2	80.8	465	1	AA443471	zw34a07.r
c 135	20.2	80.8	476	1	A1371732	tal1h12.x
c 136	20.2	80.8	481	1	AA779824	af45d12.s
c 137	20.2	80.8	494	9	B32833	HS-1016-A1-
c 138	20.2	80.8	500	5	BX508399	DKFZP686G
c 139	20.2	80.8	503	3	BQ188952	UI-E-EJ1-
c 140	20.2	80.8	505	1	AI963221	wc29a04.x
c 141	20.2	80.8	509	1	AA570379	nk62b02.s
c 142	20.2	80.8	510	9	AZ095298	RPCI-23-1
c 143	20.2	80.8	510	9	AZ607988	IM0430B18
c 144	20.2	80.8	512	2	BF287556	EST452147
c 145	20.2	80.8	516	3	BM564969	ij71c03.y
c 146	20.2	80.8	519	1	AV718629	AV718629
c 147	20.2	80.8	525	2	BF821895	WRI-RT004
c 148	20.2	80.8	530	1	AA829944	od40g12.s
c 149	20.2	80.8	535	9	AQ458762	HS_5060_B
c 150	20.2	80.8	535	9	AQ672371	HS_2276_B

ALIGNMENTS

RESULT 1	BB056745	246 bp	mRNA	linear	EST 01-AUG-2000
LOCUS	BB056745	RIKEN full-length enriched, 2 days neonate sympathetic ganglion	Mus musculus	cDNA clone 7120446E16 3', mRNA sequence.	
DEFINITION	BB056745				
ACCESSION	BB056745.1	GI:8463893			
VERSION	Est.				
KEYWORDS	Mus musculus (house mouse)				
SOURCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Mus.				
ORGANISM	Mus musculus				

REFERENCE AUTHORS

1 (bases 1 to 246)
Konno,H., Aizawa,K., Akahira,S., Akiyama,J., Arakawa,T., Carninci,P., Endo,T., Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N., Hirozane,T., Hori,F., Ishii,Y., Ishikawa,J., Ishikawa,T., Itoh,M., Izawa,M., Kadota,K., Kagawa,I., Kai,C., Kawai,J., Kikuchi,N., Kiyosawa,H., Kojima,Y., Kondo,S., Koya,S., Kurihara,C., Kusakabe,M., Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H., Okazaki,Y., Ono,T., Owa,C., Saito,H., Sakai,C., Sato,K., Shibata,K., Shibata,Y., Shigenoto,Y., Shinagawa,A., Shiraki,T., Soabe,Y., Suganara,Y., Suzuki,H., Suzuki,H., Tagawa,A., Takahashi,F., Tominaga,N., Toyota,T., Tsunoda,Y., Watanabe,A., Watanabe,S., Yamamura,T., Yamanaka,I., Yano,R., Yasunishi,A., Yokota,T., Yoshida,K., Yoshiki,A., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
RIKEN Mouse ESTs (Konno,H., et al.)
Unpublished (2000)
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Sushiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-resesc.riken.jp, URL:http://genome.gsc.riken.jp/
Carninci,P., Nishiyama,Y., Westover,A., Itoh,M., Nagaoaka,S., Sasaki,N., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Thermotabilization and thermoactivation of thermolabile enzymes by trehalose and its application for the synthesis of full length cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
Itoh,M., Kitsuina,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J., Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M., Okazaki,Y. and Hayashizaki,Y.
Automated filtration-based high-throughput plasmid preparation system. Genome Res. 9 (5), 463-470 (1999)
Carninci,P. and Hayashizaki,Y.
High-efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)
Please visit our web site (http://genome.rtc.riken.go.jp) for further details.
Location/Qualifiers
1..246
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="7120446E16"
/sex="mixed"
/tissue_type="sympathetic ganglion"
/dev_stage="2 days neonate"
/lab_host="DH10B"
/clone_lib="RIKEN full-length enriched, 2 days neonate sympathetic ganglion"
/note="Site 1: SalI; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGAGAGAGAGATCCAGAGCTCTTTTCTTTTCTTTTCTT 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGATCTCGATTAAATTAATTCCTCCCTCCCTCCCTCC 3']. cDNA was cleaved with XhoI and BamHI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC 1. Cloning sites, 5' end; SalI; 3' end; BamHI. Host: DH10B. -RNA was provided by Akira Nakagawara, Div. of Biochemistry, Chiba Cancer Center Research Institute, 666-2 Nitona, Chuoh-ku, Chiba, 260-8717 Japan, whose assistance we gratefully acknowledge."

FEATURES source

ORIGIN

```

Query Match      100.0%; Score 25; DB 1; Length 246;
Best Local Similarity 100.0%; Pred. No. 4.7;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAAAAAAAAATCAGGCCAGGCATGG 25
    |||||||
Db 100 AAAAAAAAAATCAGGCCAGGCATGG 124

RESULT 2
AW413970/c
LOCUS      AW413970.1 x1 NCI_CGAP_Mam3 Mus musculus cDNA clone IMAGE:2650485 3',
DEFINITION mRNA sequence.
ACCESSION  AW413970
VERSION     AW413970.1 GI:6940118
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 329)
            NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
AUTHORS     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
TITLE       Tumor Gene Index
JOURNAL     Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs@mail.nih.gov
            Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
            CDNA Library Preparation: Life Technologies, Inc.
            CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

MGI:1030937
Seq primer: -40UP from Gibco
High quality sequence stop: 304.
Location/Qualifiers
FEATURES             source
    source
    1..329
        /organism="Mus musculus"
        /mol_type="mRNA"
        /strain="129,C57BL/6J,FVB/N"
        /db_xref="taxon:10090"
        /clone="IMAGE:2650485"
        /tissue_type="tumor, gross tissue"
        /dev_stage="10 months"
        /lab_hosts="DH10B"
        /clone_lib="NCI_CGAP Mam3"
        /note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;
        Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
        Library constructed by Life Technologies. Investigators
        providing samples: Lothar Hennighausen/Chu-Xia Deng, NIH
        Reference for transgenic model: Xu et al., Nature Genetics
        22, 37-43 (1999)."
```

```

mRNA sequence.
AW763628
AW763628.1 GI:7695569
EST.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.
1 (bases 1 to 410)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Other ESTs: ur63e02.y1
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
image.llnl.gov/image/html/iresources.shtml

MGI:1057726
Seq primer: -40UP from Gibco
High quality sequence stop: 383.
Location/Qualifiers
FEATURES             source
    source
    1..410
        /organism="Mus musculus"
        /mol_type="mRNA"
        /strain="129,C57BL/6J,FVB/N"
        /db_xref="taxon:10090"
        /clone="IMAGE:3154970"
        /tissue_type="tumor, gross tissue"
        /dev_stage="10 months"
        /lab_hosts="DH10B"
        /clone_lib="NCI_CGAP Mam3"
        /note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;
        Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
        Library constructed by Life Technologies. Investigators
        providing samples: Lothar Hennighausen/Chu-Xia Deng, NIH
        Reference for transgenic model: Xu et al., Nature Genetics
        22, 37-43 (1999)."
```

ORIGIN

```

Query Match      100.0%; Score 25; DB 1; Length 410;
Best Local Similarity 100.0%; Pred. No. 4.9;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 AAAAAAAAAATCAGGCCAGGCATGG 25
    |||||||
Db 151 AAAAAAAAAATCAGGCCAGGCATGG 127

```

RESULT 4

```

CKX43886/c
LOCUS      CKX43886.3
DEFINITION K0806E06-3 NIA Mouse 8.5-dpc whole Embryo CDNA Library (long) Mus
            musculus cDNA clone NIA:K0806E06 IMAGE:30078965 3', mRNA sequence.
ACCESSION  CKX43886
VERSION     CKX43886.1 GI:40299499
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
1 (bases 1 to 469)
Piao.Y., Ko.N.T., Lim.M.K. and Ko.M.S.H.
Construction of long-transcript enriched cDNA libraries from
submicrogram amounts of total RNAs by a universal PCR amplification

```

JOURNAL PUBMED COMMENT	method		REFERENCE AUTHORS	
	Genome Res. 11 (9), 1553-1558 (2001) 11544199 Contact: Dawood B. Dudekula Laboratory of Genetics National Institute on Aging/National Institutes of Health 333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA Email: cdna@lgsun.grc.nia.nih.gov Plate: K0806 row: E column: 06 Seq primer: -21M13 Forward High quality sequence stop: 469 POLYA=Yes.			
FEATURES source	Location/Qualifiers		TITLE JOURNAL COMMENT	
	1. .469 /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6J" /db_xref="niaEST:K0806E06-3" /db_xref="taxon:10090" /clone="NIA:K0806E06 IMAGE:30078965" /tissue_type="whole embryo including extraembryonic tissues at 8.5-days postcoitum" /dev_stage="8.5-days postcoitum" /lab_host="DH10B" /clone_lib="NIA Mouse 8.5-dpc Whole Embryo cDNA Library (Long)" /notes="Vector: pSPORT1 (Invitrogen); Site 1: SalI; Site 2: NotI; Mouse cDNA project by the Laboratory of Genetics, National Institute on Aging (NIA), Intramural Research Program, NIH (http://lgsun.grc.nia.nih.gov/cDNA). This is a long-transcript enriched cDNA library (Ref. Genome Res. 11: 1553-1558 (2001). [PMID: 11544199]). Total RNAs were extracted from a pool of 13 embryos at 8.5-days postcoitum. Double-stranded cDNAs were synthesized with an Oligo(dT) primer [Invitrogen]: 5'-pGACTAGTCTAGATCGGAGCGCGCCCTTTT-3' from 9.1 ug of total RNA, treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to Loner-linker LL-Sal4, purified by phenol/chloroform, and separated from free linkers by Centricon 100. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer Sal4-S. The products were purified by phenol/chloroform and Centricon 100. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pSPORT1 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by the standard chemical method. The average insert size is about 2.5 kb. The library was constructed by Yulan Piao (NIA)."			
ORIGIN				
Query Match		100.0%; Score 25; DB 7; Length 469;		
Best Local Similarity		100.0%; Pred. No. 5;		
Matches 25; Conservative		0; Mismatches 0; Indels 0; Gaps 0;		
Qy	1 AAAAAAAAAATCAGGCAGGCATGG 25			
Db	 147 AAAAAAAAAATCAGGCAGGCATGG 123			
RESULT 5				
BB693815				
LOCUS	534 bp mRNA linear EST 10-OCT-2001			
DEFINITION	BB693815 RIKEN full-length enriched, 2 days neonate sympathetic ganglion Mus musculus cDNA clone 7120449D03 3', mRNA sequence.			
BB693815				
ACCESSION				
VERSION	BB693815.1 GI:16020548			
KEYWORDS	EST.			
SOURCE	Mus musculus (house mouse)			
ORGANISM	Mus musculus			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Mus.			

1 (bases 1 to 534)
Akimura,T., Arakawa,T., Carninci,P., Furuno,M., Hanagaki,T.,
Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Imotani,K.,
Ishii,Y., Ito,M., Kawai,J., Kojima,Y., Konno,H., Kouda,M.,
Matsuyama,T., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R.,
Okazaki,Y., Okido,T., Saito,K., Sakai,C., Sakai,K., Sakazume,N.,
Sasagi,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Suzuki,H., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
Tanaka,T., Tomaru,A., Toya,T., Watahiki,A., Yasunishi,A.,
Muramatsu,M. and Hayashizaki,Y.
RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al.
2001)
Unpublished (2001)
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalize and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
wagi,K., Fujiwaki,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Kira,A.,
Matsura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A.
and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P.,
Sugahara,Y. and Hayashizaki,Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Please visit our web site (http://genome.gsc.riken.go.jp) for
further details.
e mouse tissues.
Location/Qualifiers
1. .534
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="7120449D03"
/sex="mixed"
/tissue_type="sympathetic ganglion"
/dev_stage="2 days neonate"
/lab_host="DH10B"
/clone_lib="RIKEN full-length enriched, 2 days neonate
sympathetic ganglion"
/note="Site 1: SalI; Site 2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5',
GAGAGAGAGAGATCCAGAGCTCTTTTCTTTTCTTTT 3'], cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. Second strand cDNA was prepared with the
primer adapter of sequence [5',
GAGAGAGATTCGAGTTAATTAATTAATCCCCCCCCCC 3']. cDNA
was cleaved with XhoI and BamHI. Vector: a modified
pBluescript KS(+) after bulk excision from Lambda FLC I.
Cloning sites, 5' end: SalI; 3' end: BamHI. Host: DH10B.
-RNA was provided by Akira Nakagawara, Div. of
Biochemistry, Chiba Cancer Center Research Institute,
666-2 Nitona, Chuo-ku, Chiba, 260-8717 Japan, whose

assistance we gratefully acknowledge."

ORIGIN
Query Match 100.0%; Score 25; DB 2; Length 534;
Best Local Similarity 100.0%; Pred. No. 5;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAAAAAAATCAGCGCCAGGCATGG 25
|||||
Db 388 AAAAAAAAAATCAGCGCCAGGCATGG 412
|||||

RESULT 6
CB248873 649 bp mRNA linear EST 15-JUL-2003
LOCUS UI-M-EX0-byj-d-09-0-UI.r1 NIH_BMAP_EX0 Mus musculus cDNA clone
DEFINITION IMAGE:57118872 5', mRNA sequence.
CB248873
ACCESSION CB248873.1 GI:28387712
VERSION EST.
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus

ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 649)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-x@mail.nih.gov

Tissue Procurement: Dr. James Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>

This clone was contributed by the Brain Molecular Anatomy Project (BMAP)

The following repetitive elements were found in this cDNA

sequence: 114-169, >ID3#SINE/ID (matched complement) 195-337,

>B1 MM#SINE/Alu (matched complement) 471-512, >MT2B#LTR/MaLR

484-629, >B1 MM#SINE/Alu

Seq primer: pYX-5.

Location/Qualifiers

1. .649

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6"

/db_xref="taxon:10090"

/clone="IMAGE:57118872"

/tissue_type="whole brain"

/dev_stage="embryo 15.5 dpc"

/lab_host="DH10B (T1 phage resistant)"

/clone_lib="NIH_BMAP_EX0"

/note="Organ: brain; Vector: pYX-Asc; Site_1: EcoR I;
Site_2: Not I; The library was constructed according to
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured mRNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with an
oligo-dT primer containing a Not I site. Double stranded
cDNA was size selected according to mRNA size fraction,
ligated with EcoR I adaptor, digested with Not I, and then
cloned directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA
tail, is GTGCGTGGAA. This library was created for the
University of Iowa Mouse Brain Molecular Anatomy Project
(BMAP): 'Gene Discovery in the Developing Mouse Nervous
System', supported by National Institutes of Mental Health
(NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN

Query Match 93.6%; Score 23.4; DB 8; Length 354;
Best Local Similarity 96.0%; Pred. No. 24;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAAAAAAAATCAGCGCCAGGCATGG 25
|||||

Query Match 100.0%; Score 25; DB 6; Length 649;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAAAAAAATCAGCGCCAGGCATGG 25
|||||
Db 470 AAAAAAAAAATCAGCGCCAGGCATGG 494
|||||

RESULT 7
T69797/c 354 bp mRNA linear EST 15-MAR-1995
LOCUS yd14g04.s1 Soares fetal liver spleen INFLS Homo sapiens cDNA clone
DEFINITION IMAGE:108246 3' similar to contains Alu repetitive element; mRNA
sequence.
T69797
ACCESSION T69797.1 GI:680945
VERSION EST.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens

ORGANISM
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 354)
AUTHORS Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,
Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,
Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, P.,
Trevasakis, E., Waterston, R., Williamson, A., Wohldmann, P. and
Wilson, R.

TITLE The WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT Other ESTs: yd14g04.r1
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Insert Size: 679

High quality sequence stops: 300 Source: IMAGE Consortium, LLNL This
clone is available royalty-free through LLNL; contact the IMAGE
Consortium (info@image.llnl.gov) for further information.
Insert Length: 679 Std Error: 0.00
Seq primer: -21m13
High quality sequence stop: 300.

FEATURES

Location/Qualifiers

1. .354

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="GDB:463863"

/db_xref="taxon:9606"

/clone="IMAGE:108246"

/sex="male"

/dev_stage="20 week-post conception fetus"

/lab_host="DH10B (ampicillin resistant)"

/clone_lib="Soares fetal liver spleen INFLS"

/note="Organ: Liver and Spleen; Vector: pYX-Asc; Site_1: EcoR I;
Site_2: Eco RI;
with a modified polylinker; Site_1: Pac I; Site 2: Eco RI;
1st strand cDNA was primed with a Pac I - oligo(dT) primer
[5' AACTGGAGAAATTAATTAAGATCTTTTTTTTTTTTTTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Pac I and cloned into the Pac I
and Eco RI sites of the modified pYX3 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M.Fatima Bonaldo."

ORIGIN

Query Match 93.6%; Score 23.4; DB 8; Length 354;
Best Local Similarity 96.0%; Pred. No. 24;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAAAAAAAATCAGCGCCAGGCATGG 25
|||||

```
Db      236 AAAAAAAAAATCGCCAGGCATGG 212

RESULT 8
AA507708/c
LOCUS   linear      389 bp      mRNA      EST 20-AUG-1997
DEFINITION
ng85c03.s1 NCI CGAP Pr6 Homo sapiens cDNA clone IMAGE:941572
similar to contains Alu repetitive element; contains MER2.t3 MER2
repetitive element ; , mRNA sequence.
AA507708
AA507708.1 GI:2244147
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
1 (bases 1 to 389)
REFERENCE
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
AUTHORS
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL
Unpublished (1997)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: David G. Bostwick, M.D., Rodrigo F. Chuacui,
M.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: David B. Krizman, Ph.D.
DNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.biol.llnl.gov/bbrp/image/image.html
Insert Length: 632 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 356.

FEATURES
Location/Qualifiers
1..389
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:941572"
/sex="male"
/tissue_type="prostate"
/lab_host="DH10B"
/clone_lib="NCI CGAP Pr6"
/notes="Vector: pAMP10; mRNA made from prostatic
intraepithelial neoplasia (low-grade), cDNA made by
oligo-dT priming. Non-directionally cloned.
Size-selected on agarose gel, average insert size 600 bp.
Reference: Krizman et al. (1996) Cancer Research
56:5380-5383."
```

```
Db      236 AAAAAAAAAATCGCCAGGCATGG 212

REFERENCE
AUTHORS
Zhao, S., Nierman, W., Feldblyum, T., Malek, J., Shatsman, S., de
Akinret, B., Levins, M., McGann, S., Teegaye, G., Geer, K., Krol, M., de
Jong, P. and Fraser, C.M.
Mouse BAC End Sequences from Library RPCI-23
Unpublished (1999)
Other_GSSs: RPCI-23-375L24.TJ
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@igir.org
Clones are derived from the mouse BAC library RPCI-23. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm)
or from Resea ch Genetics (info@resgen.com). BAC end page:
http://www.tigr.org/cdb/bac\_ends/mouse/bac\_end\_intro.html
Plate: 375 row: L column: 24
Seq primer: T7
Class: BAC ends.

FEATURES
Location/Qualifiers
1..413
/organism="Mus musculus"
/mol_type="genomic DNA"
/strains="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-23-375L24"
/sex="Female"
/lab_host="DH10B"
/clone_lib="RPCI-23"
/notes="Organ: Kidney/Brain; Vector: pBACe3.6; Site 1:
EcoRI; Site 2: EcoRI; Female C57BL/6J mouse kidney and/or
brain genomic DNA was isolated and partially digested
with a combination of EcoRI and EcoRI Methylase. Size
selected DNA was cloned into the pBACe3.6 vector at the
EcoRI sites. The ligation products were transformed into
DH10B electrocompetent cells (BRL Life Technologies). "
```

```
Query Match      89.6%; Score 22.4; DB 9; Length 413;
Best Local Similarity 95.8%; Pred. No. 65;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAAAAAAAATCACGCCAGGCATG 24
|||||
Db 35 AAAAAAAAAATCACGCCAGGCAG 12
|||||

RESULT 10
AA554415
LOCUS   linear      254 bp      mRNA      linear      EST 10-JUN-2003
DEFINITION
L0240C02-3 NIA Mouse Newborn Ovary cDNA Library Mus musculus cDNA
Clone L0240C02 3', mRNA sequence.
AA554415
AA554415.2 GI:31569583
VERSION
KEYWORDS
SOURCE
ORGANISM
Mus musculus (house mouse)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 254)
REFERENCE
Tanaka, T.S., Jaradat, S.A., Lim, M.K., Kartgul, G.J., Wang, X.,
Grahovac, M.J., Pantano, S., Sano, Y., Piao, Y., Nagaraja, R., Doi, H.,
Wood, W.H. III, Becker, K.G. and Ko, M.S.H.
Genome-wide expression profiling of mid-gestation placenta and
embryo using a 15,000 mouse developmental cDNA microarray
Proc. Natl. Acad. Sci. U.S.A. 97 (16), 9127-9132 (2000)

REFERENCE
AUTHORS
TITLE
JOURNAL
```


PUBMED 10922068
COMMENT On Mar 7, 2000 this sequence version replaced gi:7199826.
Contact: George J. Kargul
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@lgaun.grc.nia.nih.gov
Plate: L0240 row: C column: 02
Seq primer: -21M13 Forward
High quality sequence stop: 254
POLYA=Yes.

8889548
Contact: Chin, H
National Institute of Mental Health
6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
20892-9643, USA
Tel: 301 443 1706
Fax: 301 443 9890
Email: mEST@mail.nih.gov
cDNA Library Preparation: M.B. Soares Lab Clone distribution:
Researchers may obtain BMAP cDNA clones from RESEARCH GENETICS. It
should be noted that Bento Soares is generating a small number of
additional specialized non-redundant arrays of BMAP cDNAs whose
availability will be considered under appropriate and limited
collaborative arrangements. The following repetitive elements were
found in this cDNA sequence: 1-82, >PB1D10#SINE/Alu 87-108,
>AT rich#Low complexity 344-364, >AT rich#Low complexity
Seq primer: M13 Reverse.

FEATURES
source

Location/Qualifiers
1. .317
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UI-M-BH3-atg-h-06-0-UI"
/dev_stage="27-32 days"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NIH BMAP M S4"
/notes="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not 1; Site 2: Eco RI; The
NIH BMAP M S4 library is a subtracted library of a series,
ultimately derived from a mixture of individually tagged
normalized libraries from ten regions of the mouse brain
(cerebellum, brain stems, olfactory bulbs, hypothalamus,
cortex, amygdala, basal ganglia, pineal gland, striatum,
hippocampus) after a series of subtractions to reduce the
representation of cDNAs from which ESTs had already been
generated. The following serially subtracted libraries
were generated in this process: NIH BMAP M S4,
NIH BMAP M S3.3, NIH BMAP M S3.2, NIH BMAP M S3.1,
NIH BMAP M S2, NIH BMAP M S1. The subtracted library
(NIH BMAP M S4) was constructed as follows: PCR amplified
cDNA inserts from NIH BMAP M S3.3, NIH BMAP M S3.2, and
NIH BMAP M S3.1 clones from which 3' ESTs had been derived
was used as a driver in a hybridization with a pool of
the NIH BMAP M S3.3, NIH BMAP M S3.2, and NIH BMAP M S3.1
libraries in the form of single-stranded circles. The
remaining single-stranded circles (subtracted library)
was purified by hydroxyapatite column chromatography,
converted to double-stranded circles and electroporated
into DH10B bacteria (Life Technologies) to generate the
NIH BMAP M S4 library. This procedure has been previously
described (Bonaldo, Lennon and Soares, Genome Research
6:791-806, 1996)"

ORIGIN

Query Match 87.2%; Score 21.8; DB 1; Length 254;
Best Local Similarity 92.0%; Pred. No. 1.1e+02;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 AAAAAAAAAATCAGCGCCAGCATGG 25
|||||||
Db 127 AAAAAAAAAATCTTGGCCAGCATGG 151
|||||||

RESULT 11

BF469089/c
LOCUS 317 bp mRNA linear EST 04-DEC-2000
DEFINITION UI-M-BH3-atg-h-06-0-UI.r1 NIH BMAP M S4 Mus musculus cDNA clone
UI-M-BH3-atg-h-06-0-UI 5', mRNA sequence.
ACCESSION BF469089
VERSION BF469089.1 GI:11538272
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Murinae; Mus.
1 (bases 1 to 317)
REFERENCE Bonaldo, M.F., Lennon, G. and Soares, M.B.
AUTHORS Normalization and subtraction: two approaches to facilitate gene
TITLE discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)

ORIGIN

Query Match 87.2%; Score 21.8; DB 2; Length 317;
Best Local Similarity 92.0%; Pred. No. 1.2e+02;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 AAAAAAAAAATCAGCGCCAGCATGG 25
|||||||
Db 47 AAAAAAAAAATCTTGGCCAGCATGG 23
|||||||

RESULT 12

BB105780/c
LOCUS 322 bp mRNA linear EST 27-JUN-2000
DEFINITION BB105780 RIKEN full-length enriched, 12 days embryo, embryonic body
between diaphragm region and neck Mus musculus cDNA clone
9430096N15.3, similar to L42293 Mus musculus acyl-coenzyme
A:cholesterol acyltransferase (ACACT) mRNA, mRNA sequence.
BB105780
ACCESSION BB105780.1 GI:8758348
VERSION BB105780.1

purified by ethanol-precipitation. The cDNAs were ligated to Lone-linker LL-Sal4, purified by phenol/chloroform, and separated from free linkers by Centricon 100. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer Sal4-S. The products were purified by phenol/chloroform and Centricon 100. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pQWV-SPORT6 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by the standard chemical method. The average insert size is about 2.7 kb. The library was constructed by Yulan Piao."

ORIGIN

Query Match 87.2%; Score 21.8; DB 7; Length 330;
Best Local Similarity 92.0%; Pred. No. 1.2e+02;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AAAAAAAAAATCAGCGCCAGGCATGG 25

Db 256 AAAAAAAAAATCTGGCCAGGCATGG 280

RESULT 14

BY690105/c
LOCUS BY690105 372 bp mRNA linear EST 16-DEC-2002
DEFINITION musculus full-length enriched, osteoclast-like cell Mus
musculus cDNA clone I420009013 3', mRNA sequence.

ACCESSION BY690105.1 GI:27097427

VERSION

KEYWORDS

SOURCE

ORGANISM

Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Murinae; Mus.

REFERENCE

AUTHORS

1 (bases 1 to 372)
Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
Nikaido, I., Osato, N., Salto, R., Suzuki, H., Yamanaka, I.,
Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,
Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D. P., Bult, C.,
Hume, D. A., Quackenbush, J., Schriml, L. M., Kanapin, A., Matsuda, H.,
Batalov, S., Beisel, K. W., Blake, J. A., Bradt, D., Brusic, V.,
Chothia, C., Corbani, L. E., Cousins, S., Dalla, E., Dragani, T. A.,
Fletcher, C. F., Forrest, A., Frazer, K. S., Gaasterland, P.,
Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S.,
Gustincich, S., Hirokawa, N., Jackson, I. J., Jarvis, E. D., Kanai, A.,
Kawaji, H., Kawasawa, Y., Kedzierski, R. M., King, B. L., Konagaya, A.,
Kurochkin, I. V., Lee, Y., Lenhard, B., Lyons, P. A., Maglott, D. R.,
Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T.,
Numata, K., Okido, T., Pavan, W. J., Pertea, G., Pesole, G.,
Petrovsky, N., Pillai, R., Pontius, J. U., Qi, D., Ramachandran, S.,
Ravasi, T., Reed, J. C., Reed, D. J., Reid, J., Ring, B. Z., Ringwald, M.,
Sandelin, A., Schneider, C., Semple, C. A., Setou, M., Shimada, K.,
Sultana, R., Takenaka, Y., Taylor, M. S., Teasdale, R. D., Tomita, M.,
Verardo, R., Wagner, L. G., Walestedt, C., Wang, Y., Watanabe, Y.,
Wells, C., Wilming, L. G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I.,
Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P.,
Hayatsu, N., Hirozawa, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,
Sakazume, N., Sato, K., Itoh, M., Hashizume, W., Imotani, K., Ishii, Y.,
Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y.,
Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K.,
Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E. S.,
Rogers, J., Birney, E., and Hayashizaki, Y.

Analysis of the mouse transcriptome based on functional annotation

of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)

12466851

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic

Sciences Center (GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-resesc.riken.jp, URL: http://genome.gsc.riken.jp/
Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S.,
Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H.,
Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R.,
Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K.,
Shiraki, T., Tagami, M., Waki, K., Watahiki, A., Muramatsu, M. and
Hayashizaki, Y. Direct Submission

Computational Analysis of Full-Length Mouse cDNAs Compared with
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)

Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)

RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA

encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

cDNA library was prepared and sequenced in Mouse Genome

Encyclopedia Project of Genome Exploration Research Group in Riken

Genomic Sciences Center and Genome Science Laboratory in RIKEN

Division of Experimental Animal Research in Riken contributed to

prepare mouse tissues.

Tissues were provided by Takashi Ishikawa (Department of Surgery

2 Yokohama City University 3-9 Fukuura, Kanazawa-ku, Yokohama

236-0004 Japan) whose assistance we gratefully acknowledge.

Please visit our web site (http://genome.gsc.riken.go.jp) for

further details.

FEATURES

source

1..372
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="I420009013"
/cell_type="osteoclast-like cell"
/clone_lib="RIKEN full-length enriched, osteoclast-like
cell"

ORIGIN

Query Match 87.2%; Score 21.8; DB 5; Length 372;
Best Local Similarity 92.0%; Pred. No. 1.2e+02;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AAAAAAAAAATCAGCGCCAGGCATGG 25

Db 124 AAAAAAAAAATCTGGCCAGGCATGG 100

RESULT 15

AW492561

LOCUS

DEFINITION

ACCSSSTON

VERSION

KEYWORDS

SOURCE

ORGANISM

Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

Sciurognathi; Muridea; Muridae; Murinae; Mus.

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

COMMENT

Genome Res. 6 (9), 791-806 (1996)

8889548

Contact: Chin, H

National Institute of Mental Health

6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD

AW492561 373 bp mRNA linear EST 24-FEB-2000

UT-M-BH3-atg-h-06-0-UI.s1 NIH BMAP M S4 Mus musculus cDNA clone

AW492561 3', mRNA sequence.

AW492561.1 GI:7062842

EST.

Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

Sciurognathi; Muridea; Muridae; Murinae; Mus.

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

COMMENT

Normalization and subtraction: two approaches to facilitate gene

discovery

Genome Res. 6 (9), 791-806 (1996)

8889548

Contact: Chin, H

National Institute of Mental Health

6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD

```

20892-9643, USA
Tel: 301 443 1706
Fax: 301 443 9890
Email: MES@mail.nih.gov

The sequence contained an oligo-dT track that was present in the
oligonucleotide that was used to prime the synthesis of first
strand cDNA and therefore this may represent a bonafide poly A
tail. The sequence tag present in the cDNA between the NotI site
and the oligo-dT track served to identify it as a clone from the
normalized brain stems library cDNA library Preparation: M.B.
Soares Lab Clone distribution: Researchers may obtain BMAP cDNA
clones from RESEARCH GENETICS. It should be noted that Bento Soares
is generating a small number of additional specialized
non-redundant arrays of BMAP cDNAs whose availability will be
considered under appropriate and limited collaborative arrangements
The following repetitive elements were found in this cDNA sequence:
256-277, >AT-rich#Low complexity 282-363, >PB1D10#SINE/Alu
Seq primer: M13 Forward
POLYA=Yes.

FEATURES             source
    Location/Qualifiers
        1..373
            /organism="Mus musculus"
            /mol_type="mRNA"
            /strain="C57BL/6J"
            /db_xref="taxon:10090"
            /clone="UI-M-BH3-atg-h-06-0-UI"
            /dev_stage="27-32 days"
            /lab_host="DH10B (Life Technologies)"
            /clone_lib="NIH BMAP M_S4"
            /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; The
NIH BMAP M_S4 library is a subtracted library of a series,
ultimately derived from a mixture of individually tagged
normalized libraries from ten regions of the mouse brain
(cerebellum, brain stems, olfactory bulbs, hypothalamus,
cortex, amygdala, basal ganglia, pineal gland, striatum,
hippocampus) after a series of subtractions to reduce the
representation of cDNAs from which ESTs had already been
generated. The following serially subtracted libraries
were generated in this process: NIH BMAP M_S4,
NIH BMAP M_S3.3, NIH BMAP M_S3.2, NIH BMAP M_S1,
NIH BMAP M_S4) was constructed as follows: PCR amplified
cDNA inserts from NIH BMAP M_S3.3, NIH BMAP M_S3.2, and
NIH BMAP M_S3.1 clones from which 3' ESTs had been derived
was used as a driver in a hybridization with a pool of
the NIH BMAP M_S3.3, NIH BMAP M_S3.2, and NIH BMAP M_S3.1
libraries in the form of single-stranded circles. The
remaining single-stranded circles (subtracted library)
was purified by hydroxyapatite column chromatography,
converted to double-stranded circles and electroporated
into DH10B bacteria (Life Technologies) to generate the
NIH BMAP M_S4 library. This procedure has been previously
described (Bonaldo, Lennon and Soares, Genome Research
6:791-805, 1996)
TAG_TISSUE=brain-stems
TAG_LIB=NIH_BMAP_M_S4
TAG_SEQ=TCATG"

ORIGIN
Query Match      87.2%; Score 21.8; DB 1; Length 373;
Best Local Similarity 92.0%; Pred. No. 1.2e+02;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAAAAAATCAGCGCCAGGCATGG 25
    |||||
Db 268 AAAAAAAAAATCTGGCCAGGCATGG 292
    |||||

RESULT 16
BB791462/c
LOCUS
DEFINITION BB791462 RIKEN full-length enriched, CRL-1751 WEHI 164 cDNA Mus

musculus cDNA clone G430136J01 3', mRNA sequence.
BB791462
BB791462.1 GI:16959958
EST.
Mus musculus (house mouse)
Mus musculus
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Murinae; Mus.
1 (bases 1 to 375)
Akimura,T., Arakawa,T., Carninci,P., Furuno,M., Hanagaki,T.,
Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Imotani,K.,
Ishii,Y., Ito,M., Kawai,J., Kojima,Y., Konno,H., Kouda,M.,
Matsuyama,T., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R.,
Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sakazume,N.,
Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Suzuki,H., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
Tanaka,T., Tomaru,A., Toya,T., Watahiki,A., Yasunishi,A.,
Muramatsu,M. and Hayashizaki,Y.
RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al.
2001)
Unpublished (2001)
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Sushiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
wagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T.,
Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A.
and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P.,
Sugahara,Y. and Hayashizaki,Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Please visit our web site (http://genome.gsc.riken.go.jp) for
further details.
e mouse tissues.
    Location/Qualifiers
        1..375
            /organism="Mus musculus"
            /mol_type="mRNA"
            /strain="BALB/c"
            /db_xref="taxon:10090"
            /clone="G430136J01"
            /cell_line="CRL-1751 WEHI 164"
            /clone_lib="RIKEN full-length enriched, CRL-1751 WEHI 164
cDNA"

Query Match      87.2%; Score 21.8; DB 2; Length 375;
Best Local Similarity 92.0%; Pred. No. 1.2e+02;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAAAAAATCAGCGCCAGGCATGG 25
    |||||
Db 121 AAAAAAAAAATCTGGCCAGGCATGG 97
    |||||

RESULT 17
BB791462/c
LOCUS
DEFINITION BB791462 RIKEN full-length enriched, CRL-1751 WEHI 164 cDNA Mus

```

LOCUS	BY698909	379 bp	mRNA	linear	EST 16-DEC-2002
DEFINITION	BY698909	RIKEN full-length enriched, osteoclast-like cell Mus musculus cDNA clone 1420113M05 3', mRNA sequence.			
ACCESSION	BY698909				
KEYWORDS	BY698909.1	GI:27109934			
SOURCE	EST.				
ORGANISM	Mus musculus (house mouse)				
REFERENCE	Mus musculus				
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridea; Muridae; Murinae; Mus.				
	1 (bases 1 to 379)				
	Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S., Nikiado,I., Osato,N., Saico,R., Suzuki,H., Yananaka,I., Kiyosawa,H., Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A., Schonbach,C., Gojobori,T., Baldarelli,R., Hill,D.P., Bult,C., Hume,D.A., Quackenbush,J., Schriml,L.M., Kanapin,A., Matsuuda,H., Batalov,S., Beisel,K.W., Blake,J.A., Bradt,D., Brusic,V., Ciothia,C., Corbani,L.E., Cousins,S., Dalla,E., Dragani,T.A., Fletcher,C.F., Forrest,A., Frazer,K.S., Gaasterland,T., Gariboldi,M., Gissi,C., Godzik,A., Gough,J., Grimmond,S., Gustinchik,S., Hirokawa,N., Jackson,I.J., Jarvis,E.D., Kanai,A., Kawaji,H., Kawasawa,Y., Kedzierski,R.M., King,B.L., Konagaya,A., Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R., Maltais,L., Marchionni,L., McKensie,L., Miki,H., Nagashima,T., Numata,K., Okido,T., Pavan,W.J., Perte,G., Pesole,G., Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramachandran,S., Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring,B.Z., Ringwald,M., Sandelin,A., Schneider,C., Sample,C.A., Setou,M., Shimada,K., Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tonita,M., Vervardo,R., Wagner,L.G., Walestedt,C., Wang,Y., Watanabe,Y., Wells,C., Wilming,L.G., Wynshaw-Boris,A., Yanagisawa,M., Yang,I., Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A., Carninci,P., Hayateu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura,M., Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K., Akazawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata,K., Shingawa,A., Yasunishi,A., Yoshino,M., Watanabe,H., Rogers,J., Birney,E. and Hayashizaki,Y.				
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs				
JOURNAL	Nature 420, 563-573 (2002)				
PUBLISHED	12466851				
COMMENT	Contact: Yoshihide Hayashizaki Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute The Institute of Physical and Chemical Research (RIKEN) 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan Tel: 81-45-503-3222 Fax: 81-45-503-9216 Email: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/ Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hirozane,T., Imotani,K., Ishii,Y., Itoh,M., Kawai,J., Konno,H., Miyazaki,A., Murata,M., Nakamura,M., Nomura,K., Numazaki,K., Ohno,M., Sakai,K., Sakazume,N., Sasaki,D., Sato,K., Shibata,K., Shiraki,T., Tagami,M., Waki,K., Watahiki,A., Muramatsu,M. and Hayashizaki,Y. Direct Submission Computational Analysis of Full-length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001) Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000) RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000) Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001) cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.				

/note="Vector: pCMV-SPORT6 (Invitrogen); Site 1: SalI; Site 2: NotI; Mouse cDNA project by the Laboratory of Genetics, National Institute on Aging (NIA), Intramural Research Program, NIH (<http://lgsun.grc.nia.nih.gov/cDNA>). This is a long-transcript enriched cDNA library (Ref. Genome Res. 11: 1553-1558 (2001). [PMID: 11544199]). Total RNAs were extracted from a pool of 3 embryos at 11.5-days postcoitum. Double-stranded cDNAs were synthesized with an Oligo(dT) primer [Invitrogen]:
 5'-pGACTAGTCTAGATCGGCGCGCCCTTTTTTTTTTTT-3' from 2ug of total RNA, treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to lone-linker LL-SalI, purified by phenol/chloroform, and separated from free linkers by Centricon 100. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer Sal4-S. The products were purified by phenol/chloroform and Centricon 100. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pCMV-SPORT6 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by the standard chemical method. The average insert size is about 3.3Kb. The library was constructed by Yulan Piao."

ORIGIN

Query Match 87.2%; Score 21.8; DB 7; Length 386;
 Best Local Similarity 92.0%; Pred. No. 1.2e+02;
 Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAAAAAATCAGCGCGCGCATGG 25

Db 255 AAAAAAAAAATCTGGCCGCGCATGG 279

RESULT 19

BY425365/C

LOCUS

DEFINITION BY425365 RIKEN full-length enriched, 16 days embryo kidney Mus

musculus cDNA clone 1920078L02 3', mRNA sequence.

BY425365

BY425365.1 GI:26699413

EST.

Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

Sciurognathi; Muroidea; Muridae; Murinae; Mus.

1 (bases 1 to 386)

Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S.,

Nikaido,I., Oatono,N., Saito,R., Suzuki,H., Yamanaka,I.,

Kiyosawa,H., Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A.,

Schonbach,C., Gojobori,T., Baldarelli,R., Hill,D.P., Bult,C.,

Hume,D.A., Quackenbush,J., Schriml,L.M., Kanapin,A., Matsuda,H.,

Batalov,S., Beisel,K.W., Blake,J.A., Bradt,D., Brusci,V.,

Chothia,C., Corbani,L.E., Cousins,S., Dalla,E., Dragani,T.A.,

Fletcher,C.F., Forrest,A., Frazer,K.S., Gaasterland,T.,

Gariboldi,M., Gissi,C., Godzik,A., Gough,J., Grimmond,S.,

Gustincich,S., Hirokawa,N., Jackson,I.J., Jarvis,E.D., Kanai,A.,

Kawaji,H., Kawasawa,Y., Kedzierski,R.M., King,B.L., Konagaya,A.,

Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R.,

Malcais,L., Marchionni,L., McKenzie,L., Miki,H., Nagashima,T.,

Numata,K., Okido,T., Pavan,W.J., Perteau,G., Pesole,G.,

Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramachandran,S.,

Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring,B.Z., Ringwald,M.,

Sandelin,A., Schneider,C., Semple,C.A., Setou,M., Shimada,K.,

Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tomita,M.,

Verardo,R., Wagner,L., Wahlestedt,C., Wang,Y., Watanabe,Y.,

Wells,C., Wilming,L.G., Wynshaw-Boris,A., Yanagisawa,M., Yang,I.,

Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A., Carninci,P.,

Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura,M.,

Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K.,

Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K., Ishii,Y.,

Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata,K.,

Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S.,

Rogers,J., Birney,E. and Hayashizaki,Y.

Analysis of the mouse transcriptome based on functional annotation

of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)

12466851

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic

Sciences Center (GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-res@gsc.riken.jp, URL:<http://genome.gsc.riken.jp/>

Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S.,

Hirozane,T., Imotani,K., Ishii,Y., Itoh,M., Kawai,J., Konno,H.,

Miyazaki,A., Murata,M., Nakamura,M., Nomura,K., Numazaki,R.,

Ohno,M., Sakai,K., Sakazume,N., Sasaki,D., Sato,K., Shibata,K.,

Shiraki,T., Tagami,M., Waki,K., Watahiki,A., Muramatsu,M. and

Hayashizaki,Y. Direct Submission

Computational Analysis of Full-length Mouse cDNAs Compared with

Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new

genes. Genome Res. 10 (10), 1617-1630 (2000)

RIKEN integrated sequence analysis (RISA) system--384-format

sequencing pipeline with 384 multicapillary sequencer. Genome Res.

10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA

encyclopedia: real-time sequence clustering for construction of a

nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

cDNA library was prepared and sequenced in Mouse Genome

Encyclopedia Project of Genome Exploration Research Group in Riken

Genomic Sciences Center and Genome Science Laboratory in RIKEN.

Division of Experimental Animal Research in Riken contributed to

prepare mouse tissues.

Please visit our web site (<http://genome.gsc.riken.go.jp>) for

further details.

Location/Qualifiers

1..388

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="1920078L02"

/tissue_type="kidney"

/dev_stage="16 days embryo"

/clone_lib="RIKEN full-length enriched, 16 days embryo

kidney"

ORIGIN

Query Match

Best Local Similarity 87.2%; Score 21.8; DB 5; Length 388;

Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAAAAAATCAGCGCGCATGG 25

Db 136 AAAAAAAAAATCTGGCGCGCATGG 112

RESULT 20

BY684912/C

LOCUS

DEFINITION BY684912 RIKEN full-length enriched, stroma cell Mus musculus cDNA

clone 1320008M16 3', mRNA sequence.

BY684912

ACCESSION BY684912

VERSION

BY684912.1 GI:27083133

KEYWORDS

Mus musculus (house mouse)

SOURCE

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

Sciurognathi; Muroidea; Muridae; Murinae; Mus.

BY684912

BY684912 RIKEN full-length enriched, stroma cell Mus musculus cDNA

clone 1320008M16 3', mRNA sequence.

BY684912

ACCESSION BY684912

VERSION

BY684912.1 GI:27083133

KEYWORDS

Mus musculus (house mouse)

SOURCE

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

Sciurognathi; Muroidea; Muridae; Murinae; Mus.

```

REFERENCE
AUTHORS
1 (bases 1 to 392)
Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S.,
Nikaido,I., Osato,N., Saito,R., Suzuki,H., Yamanaka,I.,
Kiyosawa,H., Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A.,
Schonbach,C., Gojobori,T., Baldarelli,R., Hill,D.P., Bult,C.,
Hume,D.A., Quackenbush,J., Schriml,L.M., Kanapin,A., Matsuda,H.,
Batalov,S., Beisel,K.W., Blake,J.A., Bradt,D., Bruscia,V.,
Chothia,C., Corbani,L.E., Cousins,S., Dalla,E., Dragani,T.A.,
Fletcher,C.P., Forrest,A., Frazer,K.S., Gaasterland,T.,
Gariboldi,M., Gissi,C., Godzik,A., Gough,J., Grimmond,S.,
Gustincich,S., Hirokawa,N., Jackson,I.J., Jarvis,E.D., Kanai,A.,
Kawaji,H., Kawasawa,Y., Kedzierski,R.M., King,B.L., Konagaya,A.,
Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R.,
Maltais,L., Marchionni,L., McKenzie,L., Miki,H., Nagashima,T.,
Numata,K., Okido,T., Pavan,W.J., Pertea,G., Pesole,G.,
Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramachandran,S.,
Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring,B.Z., Ringwald,M.,
Sandelin,A., Schneider,C., Semple,C.A., Setou,M., Shimada,K.,
Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tomita,M.,
Verardo,R., Wagner,L., Wahlestedt,C., Wang,Y., Watanabe,Y.,
Wells,C., Wilming,L.G., Wynshaw-Boris,A., Yanagisawa,M., Yang,I.,
Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A., Carninci,P.,
Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura,M.,
Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K.,
Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K., Ishii,Y.,
Itch,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata,K.,
Shingawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S.,
Rogers,J., Birney,E. and Hayashizaki,Y.

TITLE
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs

JOURNAL
NATURE 420, 563-573 (2002)
PUBMED
12466851
COMMENT
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gs.riken.jp, URL: http://genome.gsc.riken.jp/
Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S.,
Hirozane,T., Imotani,K., Ishii,Y., Itoh,M., Kawai,J., Konno,H.,
Miyazaki,A., Murata,M., Nakamura,M., Nomura,K., Numazaki,R.,
Ohno,M., Sakai,K., Sakazume,N., Sasaki,D., Sato,K., Shibata,K.,
Shiraki,T., Tagami,M., Waki,K., Watahiki,A., Muramatsu,M. and
Hayashizaki,Y. Direct Submission
Computational Analysis of Full-Length Mouse cDNAs Compared with
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Tissues were provided by Takashi Ishikawa (Department of Surgery
2 Yokohama City University 3-9 Fukuura, Kanazawa-ku, Yokohama
236-0004 Japan) whose assistance we gratefully acknowledge.
Please visit our web site (http://genome.gsc.riken.go.jp) for
further details
Location/Qualifiers
1. .392
/organisms="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"

FEATURES
Source
/clone="I320008M16"
/cell_type="stroma cell"
/clone_lib="RIKEN full-length enriched, stroma cell"

ORIGIN
Query Match 87.2%; Score 21.8; DB 5; Length 392;
Best Local Similarity 92.0%; Pred. No. 1.2e+02;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAAAAAATCAGCGCAGCGCATGG 25
|||||
Db 134 AAAAAAAAAATCTGGCCAGCGCATGG 110

RESULT 21
BY691014/c
LOCUS
BY691014 RIKEN full-length enriched, osteoclast-like cell Mus
DEFINITION
musculus cDNA clone 1420015L11 3', mRNA sequence.
ACCESSION
BY691014
VERSION
BY691014.1 GI:27099180
KEYWORDS
Mus musculus (house mouse)
SOURCE
Mus musculus
ORGANISM
Mus musculus
REFERENCE
Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S.,
Nikaido,I., Osato,N., Saito,R., Suzuki,H., Yamanaka,I.,
Kiyosawa,H., Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A.,
Schonbach,C., Gojobori,T., Baldarelli,R., Hill,D.P., Bult,C.,
Hume,D.A., Quackenbush,J., Schriml,L.M., Kanapin,A., Matsuda,H.,
Batalov,S., Beisel,K.W., Blake,J.A., Bradt,D., Bruscia,V.,
Chothia,C., Corbani,L.E., Cousins,S., Dalla,E., Dragani,T.A.,
Fletcher,C.P., Forrest,A., Frazer,K.S., Gaasterland,T.,
Gariboldi,M., Gissi,C., Godzik,A., Gough,J., Grimmond,S.,
Gustincich,S., Hirokawa,N., Jackson,I.J., Jarvis,E.D., Kanai,A.,
Kawaji,H., Kawasawa,Y., Kedzierski,R.M., King,B.L., Konagaya,A.,
Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R.,
Maltais,L., Marchionni,L., McKenzie,L., Miki,H., Nagashima,T.,
Numata,K., Okido,T., Pavan,W.J., Pertea,G., Pesole,G.,
Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramachandran,S.,
Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring,B.Z., Ringwald,M.,
Sandelin,A., Schneider,C., Semple,C.A., Setou,M., Shimada,K.,
Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tomita,M.,
Verardo,R., Wagner,L., Wahlestedt,C., Wang,Y., Watanabe,Y.,
Wells,C., Wilming,L.G., Wynshaw-Boris,A., Yanagisawa,M., Yang,I.,
Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A., Carninci,P.,
Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura,M.,
Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K.,
Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K., Ishii,Y.,
Itch,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata,K.,
Shingawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S.,
Rogers,J., Birney,E. and Hayashizaki,Y.

TITLE
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs

JOURNAL
NATURE 420, 563-573 (2002)
PUBMED
12466851
COMMENT
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gs.riken.jp, URL: http://genome.gsc.riken.jp/
Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S.,
Hirozane,T., Imotani,K., Ishii,Y., Itoh,M., Kawai,J., Konno,H.,
Miyazaki,A., Murata,M., Nakamura,M., Nomura,K., Numazaki,R.,
Ohno,M., Sakai,K., Sakazume,N., Sasaki,D., Sato,K., Shibata,K.,
Shiraki,T., Tagami,M., Waki,K., Watahiki,A., Muramatsu,M. and
Hayashizaki,Y. Direct Submission
Computational Analysis of Full-Length Mouse cDNAs Compared with
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Tissues were provided by Takashi Ishikawa (Department of Surgery
2 Yokohama City University 3-9 Fukuura, Kanazawa-ku, Yokohama
236-0004 Japan) whose assistance we gratefully acknowledge.
Please visit our web site (http://genome.gsc.riken.go.jp) for
further details
Location/Qualifiers
1. .392
/organisms="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"

```


Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
 Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
 RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)
 Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.
 Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
 Tissues were provided by Takashi Ishikawa (Department of Surgery 2 Yokohama City University 3-9 Fukuura, Kanazawa-ku, Yokohama 236-0004 Japan) whose assistance we gratefully acknowledge.
 Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES

source

```

1. .394
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="I420015L11"
/cell_type="osteoclast-like cell"
/clone_lib="RIKEN full-length enriched, osteoclast-like cell"

```

ORIGIN

```

Query Match      87.2%; Score 21.8; DB 5; Length 394;
Best Local Similarity 92.0%; Pred. No. 1.2e+02;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 1 AAAAAAAAAATCAGCGCCAGGCATGG 25

```

|||||
Db 154 AAAAAAAAAATCTGGCCAGGCATGG 130

```

RESULT 22

BY701163/c

```

LOCUS
DEFINITION
Mus musculus full-length enriched, osteoclast-like cell Mus
musculus cDNA clone I420130B21 3', mRNA sequence.

```

BY701163

BY701163.1 GI:27112232

EST.

Mus musculus (house mouse)

Mus musculus

ORGANISM

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Murinae; Mus.

```

1 (bases 1 to 401)

```

Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S.,
Nikaido,I., Oatso,N., Saito,R., Suzuki,H., Yamanaka,I.,
Kiyosawa,H., Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A.,
Schonbach,C., Gojobori,T., Baldarelli,R., Hill,D.P., Bult,C.,
Hume,D.A., Quackenbush,J., Schriml,L.M., Kanapin,A., Matsuda,H.,
Batalov,S., Beisel,K.W., Blake,J.A., Bradt,D., Brusic,V.,
Chothia,C., Corbani,L.E., Cousins,S., Dalla,E., Dragani,T.A.,
Fletcher,C.F., Forrest,A., Frazer,K.S., Gaasterland,T.,
Gariboldi,M., Gissi,C., Godzik,A., Gough,J., Grimmond,S.,
Gustincich,S., Hirokawa,N., Jackson,I.J., Jarvis,E.D., Kanai,A.,
Kawaji,H., Kawasawa,Y., Kedzierski,R.M., King,B.L., Konagaya,A.,
Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R.,
Maltais,L., Marchionni,L., McKenzie,L., Miki,H., Nagashima,T.,
Numata,K., Okido,T., Pavan,W.J., Pertea,G., Pesole,G.,
Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramachandran,S.,
Ravasi,T., Read,J.C., Read,D.J., Reid,J., Ring,B.Z., Ringwald,M.,
Sandelin,A., Schneider,C., Sempie,C.A., Setou,M., Shimada,K.,

```

Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tomita,M.,
 Verardo,R., Wagner,L., Wahlestedt,C., Wang,Y., Watanabe,Y.,
 Wells,C., Wilming,L.G., Wynshaw-Boris,A., Yanagisawa,M., Yang,I.,
 Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A., Carninci,P.,
 Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura,M.,
 Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K.,
 Akazawa,T., Fukuda,S., Hata,A., Hashizume,W., Imotani,K., Ishii,Y.,
 Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,K., Shibata,K.,
 Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S.,
 Rogers,J., Birney,E. and Hayashizaki,Y.
 Analysis of the mouse transcriptome based on functional annotation
 of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)
 12466851
 Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@gsc.riken.jp, URL:<http://genome.gsc.riken.jp/>
 Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S.,
 Hirozane,T., Imotani,K., Ishii,Y., Itoh,M., Kawai,J., Konno,H.,
 Miyazaki,A., Murata,M., Nakamura,M., Nomura,K., Numazaki,R.,
 Ohno,M., Sakai,K., Sakazume,N., Sasaki,D., Sato,K., Shibata,K.,
 Shiraki,T., Tagami,M., Waki,K., Watahiki,A., Muramatsu,M. and
 Hayashizaki,Y. Direct Submission
 Computational Analysis of Full-Length Mouse cDNAs Compared with
 Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new
 genes. Genome Res. 10 (10), 1617-1630 (2000)
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.
 10 (11), 1757-1771 (2000)
 Computer-based methods for the mouse full-length cDNA
 encyclopedia: real-time sequence clustering for construction of a
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 cDNA library was prepared and sequenced in Mouse Genome
 Encyclopedia Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.
 Division of Experimental Animal Research in Riken contributed to
 prepare mouse tissues.
 Tissues were provided by Takashi Ishikawa (Department of Surgery
 2 Yokohama City University 3-9 Fukuura, Kanazawa-ku, Yokohama
 236-0004 Japan) whose assistance we gratefully acknowledge.
 Please visit our web site (<http://genome.gsc.riken.go.jp>) for
 further details.

TITLE

JOURNAL

PUBMED

COMMENT

FEATURES

source

```

1. .401
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="I420130B21"
/cell_type="osteoclast-like cell"
/clone_lib="RIKEN full-length enriched, osteoclast-like cell"

```

ORIGIN

```

Query Match      87.2%; Score 21.8; DB 5; Length 401;
Best Local Similarity 92.0%; Pred. No. 1.2e+02;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 1 AAAAAAAAAATCAGCGCCAGGCATGG 25

```

|||||
Db 141 AAAAAAAAAATCTGGCCAGGCATGG 117

```

RESULT 23

BY694371/c

LOCUS

```

BY694371
403 bp mRNA linear EST 16-DEC-2002

```


DEFINITION	BY694371 RIKEN full-length enriched, osteoclast-like cell Mus musculus cDNA clone 1420033A09 3', mRNA sequence.
ACCESSION	BY694371
VERSION	BY694371.1 GI:27105326
KEYWORDS	EST.
SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus
REFERENCE	Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridea; Muridae; Murinae; Mus.
AUTHORS	1 (bases 1 to 403) Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S., Nikaido,I., Osato,N., Saito,R., Suzuki,H., Yamanaka,I., Kiyosawa,H., Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A., Schonbach,C., Gojobori,T., Baldarelli,R., Hill,D.P., Bult,C., Hume,D.A., Quackenbush,J., Schriml,L.M., Kanapin,A., Matsuda,H., Batalov,S., Beisel,K.W., Blake,J.A., Bradt,D., Brusic,V., Chothia,C., Corbani,L.E., Cousins,S., Dalla,E., Dragani,T.A., Fletcher,C.F., Forrest,A., Frazer,K.S., Gaasterland,T., Gariboldi,M., Gissi,C., Godzik,A., Gough,J., Grimmond,S., Gustincich,S., Hirokawa,N., Jackson,I.J., Jarvis,E.D., Kanai,A., Kawaji,H., Kawasaki,Y., Kedzierzki,R.M., King,B.L., Konagaya,A., Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R., Maltais,L., Marchionni,L., McKenzie,L., Miki,H., Nagashima,T., Numata,K., Okido,T., Pavan,W.J., Pertea,G., Pesole,G., Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramachandran,S., Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tomita,M., Verardo,R., Wagner,L., Wahlestedt,C., Wang,Y., Watanabe,Y., Wells,C., Wilming,L.G., Wynshaw-Boris,A., Yanagisawa,M., Yang,I., Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A., Carninci,P., Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura,M., Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K., Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata,K., Shingawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S., Rogers,J., Birney,E. and Hayashizaki,Y.
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL	Nature 420, 563-573 (2002)
PUBLISHED	12466851
COMMENT	Contact: Yoshihide Hayashizaki Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan Tel: 81-45-503-9222 Fax: 81-45-503-9216 Email: genome-res@gscc.riken.jp, URL: http://genome.gsc.riken.jp/ Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hirozane,T., Imotani,K., Ishii,Y., Itoh,M., Kawai,J., Konno,H., Miyazaki,A., Murata,M., Nakamura,M., Nomura,K., Numazaki,R., Ohno,M., Sakai,K., Sakazume,N., Sasaki,D., Sato,K., Shibata,K., Shiraki,T., Tagami,M., Waki,K., Watahiki,A., Muramatsu,M. and Hayashizaki,Y. Direct Submission Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001) Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000) RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000) Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001) cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Tissues were provided by Takashi Ishikawa (Department of Surgery
FEATURES	source Location/Qualifiers 1..403 /organism="Mus musculus" /mol_type="mRNA" /strain="CS7BL/6J" /db_xref="taxon:10090" /clone="1420033A09" /cell_type="osteoclast-like cell" /clone_lib="RIKEN full-length enriched, osteoclast-like cell"
ORIGIN	Query Match 87.2%; Score 21.8; DB 5; Length 403; Best Local Similarity 92.0%; Pred. No. 1.2e+02; Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0; QY 1 AAAAAAAAAATCAGCGCAGCGCATGG 25 Db 146 AAAAAAAAAATCTGGCCAGCGCATGG 122 RESULT 24 BY690774/C LOCUS DEFINITION ACCESSION VERSION KEYWORDS SOURCE ORGANISM REFERENCE AUTHORS Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S., Nikaido,I., Osato,N., Saito,R., Suzuki,H., Yamanaka,I., Kiyosawa,H., Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A., Schonbach,C., Gojobori,T., Baldarelli,R., Hill,D.P., Bult,C., Hume,D.A., Quackenbush,J., Schriml,L.M., Kanapin,A., Matsuda,H., Batalov,S., Beisel,K.W., Blake,J.A., Bradt,D., Brusic,V., Chothia,C., Corbani,L.E., Cousins,S., Dalla,E., Dragani,T.A., Fletcher,C.F., Forrest,A., Frazer,K.S., Gaasterland,T., Gariboldi,M., Gissi,C., Godzik,A., Gough,J., Grimmond,S., Gustincich,S., Hirokawa,N., Jackson,I.J., Jarvis,E.D., Kanai,A., Kawaji,H., Kawasaki,Y., Kedzierzki,R.M., King,B.L., Konagaya,A., Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R., Maltais,L., Marchionni,L., McKenzie,L., Miki,H., Nagashima,T., Numata,K., Okido,T., Pavan,W.J., Pertea,G., Pesole,G., Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramachandran,S., Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tomita,M., Verardo,R., Wagner,L., Wahlestedt,C., Wang,Y., Watanabe,Y., Wells,C., Wilming,L.G., Wynshaw-Boris,A., Yanagisawa,M., Yang,I., Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A., Carninci,P., Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura,M., Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K., Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata,K., Shingawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S., Rogers,J., Birney,E. and Hayashizaki,Y.

The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216

Email: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/
Aizawa.K., Akimura.T., Arakawa.T., Carninci.P., Fukuda.S.,
Hirozane.T., Imotani.K., Ishii.Y., Itoh.M., Kawai.J., Konno.H.,
Miyazaki.A., Murata.M., Nakamura.M., Nomura.K., Numazaki.R.,
Ohno.M., Sakai.K., Sakazume.N., Sasaki.D., Sato.K., Shibata.K.,
Shiraki.T., Tagami.M., Waki.K., Watahiki.A., Muramatsu.M. and
Hayashizaki.Y. Direct Submission
Computational Analysis of Full-Length Mouse cDNAs Compared with
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Tissues were provided by Takashi Ishikawa (Department of Surgery
2 Yokohama City University 3-9 Fukuura, Kanazawa-ku, Yokohama
236-0004 Japan) whose assistance we gratefully acknowledge.
Please visit our web site (http://genome.gsc.riken.go.jp) for
further details.

FEATURES source
1. .404
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="I420014G11"
/cell_type="osteoclast-like cell"
/clone_lib="RIKEN full-length enriched, osteoclast-like
cell"

ORIGIN
Query Match 87.2%; Score 21.8; DB 5; Length 404;
Best Local Similarity 92.0%; Pred. No. 1.2e+02;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 AAAAAAAAAATCAGGCCAGGCATGG 25
145 AAAAAAAAAATCTGGCCAGGCATGG 121
|||||||
|||||||

RESULT 25
BY655013/c
LOCUS
DEFINITION
cDNA clone K530321K23 3', mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Murinae; Mus.
1 (bases 1 to 415)
Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S.,
Nikaido,I., Osato,N., Saito,R., Suzuki,H., Yamanaka,I.,
Kiyosawa,H., Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A.,
Schonbach,C., Gojobori,T., Baldarelli,R., Hill,D.P., Bult,C.,
Hume,D.A., Quackenbush,J., Schriml,L.M., Kanapin,A., Matsuda,H.,
Batalov,S., Beisel,K.W., Blake,J.A., Bradt,D., Bruaic,V.,

Chothia,C., Corbani,L.E., Cousins,S., Dalla,E., Dragani,T.A.,
Fletcher,C.F., Forrest,A., Frazer,K.S., Gaasterland,T.,
Gariboldi,M., Gissi,C., Godzik,A., Gough,J., Grimmond,S.,
Gustincich,S., Hirokawa,N., Jackson,I.J., Jarvis,E.D., Kanai,A.,
Kawaji,H., Kawasawa,Y., Kedzierski,R.M., King,B.L., Kongaya,A.,
Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R.,
Maltais,L., Marchionni,L., McKenzie,L., Miki,H., Nagashima,T.,
Numata,K., Okido,T., Pavan,W.J., Pertea,G., Pesole,G.,
Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramchandran,S.,
Ravasi,I., Reed,J.C., Reed,D.J., Reid,J., King,B.Z., Ringwald,M.,
Sandelin,A., Schneider,C., Semple,C.A., Setou,M., Shimada,K.,
Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tomita,M.,
Vetardo,R., Wagner,L., Wahlestedt,C., Wang,Y., Watanabe,Y.,
Wells,C., Wilming,L.G., Wynshaw-Boris,A., Yanagisawa,M., Yang,I.,
Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A., Carninci,P.,
Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura,M.,
Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K.,
Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K., Ishii,Y.,
Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata,K.,
Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S.,
Rogers,J., Birney,E. and Hayashizaki,Y.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
12466851
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/
Aizawa.K., Akimura.T., Arakawa.T., Carninci.P., Fukuda.S.,
Hirozane.T., Imotani.K., Ishii.Y., Itoh.M., Kawai.J., Konno.H.,
Miyazaki.A., Murata.M., Nakamura.M., Nomura.K., Numazaki.R.,
Ohno.M., Sakai.K., Sakazume.N., Sasaki,D., Sato,K., Shibata,K.,
Shiraki,T., Tagami.M., Waki,K., Watahiki,A., Muramatsu,M. and
Hayashizaki,Y. Direct Submission
Computational Analysis of Full-length Mouse cDNAs Compared with
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Tissues were provided by Michela Fagiolini and Takao K. Hensch (
Laboratory for Neuronal Circuit Development Brain Science Institute
RIKEN 2-1 Hirosawa,Wako-shi,Saitama 351-0198 Japan) whose
assistance we gratefully acknowledge.
Please visit our web site (http://genome.gsc.riken.go.jp) for
further details.

FEATURES source
1. .415
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="K530321K23"
/tissue_type="visual cortex"
/clone_lib="RIKEN full-length enriched, visual cortex"

ORIGIN
Query Match 87.2%; Score 21.8; DB 5; Length 415;


```

/strain="C57BL/6J"
/db_xref="taxon:10090"
/clon="I420102C01"
/cell_type="osteoclast-like cell"
/clon_lib="RIKEN full-length enriched, osteoclast-like
cell"

ORIGIN

Query Match      87.2%; Score 21.8; DB 5; Length 423;
Best Local Similarity 92.0%; Pred. No. 1.2e+02;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAAAAAATCAGCGCCAGGCATGG 25
    |||||
Db 176 AAAAAAAAAATCTGGCCAGGCATGG 152

RESULT 30
BB781154 425 bp mRNA linear EST 08-JUL-2003
LOCUS
DEFINITION BB781154 RIKEN full-length enriched, B cells CRL-1669 BCL1 Clone
13.20-3B3 cDNA Mus musculus cDNA clone G430071G05 3', mRNA
sequence.
BB781154
BB781154.1 GI:16941854
EST.
Mus musculus (house mouse)
SOURCE
ORGANISM
Mus musculus

REFERENCE
AUTHORS
Akimura,T., Arakawa,T., Carninci,P., Furuno,M., Hanagaki,T.,
Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Imotani,K.,
Iehii,Y., Ito,M., Kawai,J., Kojima,Y., Konno,H., Kouda,M.,
Matsuyama,T., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R.,
Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sakazume,N.,
Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Suzuki,H., Tagawa,A., Tanaka,T., Tanaka,T., Tanaka,T.,
Tanaka,T., Tomaru,A., Toyota,T., Watahiki,A., Yasunishi,A.,
Muramatsu,M. and Hayashizaki,Y.
RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al.
2001)

TITLE
Unpublished (2001)
JOURNAL
COMMENT
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. . 10 (10), 1617-1630 (2000)
wagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T.,
Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A.
and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
. 10 (11), 1757-1771 (2000)
Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P.,
Sugahara,Y. and Hayashizaki,Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. . 11 (2), 281-289 (2001)
Please visit our web site (http://genome.gsc.riken.go.jp) for
further details.
e mouse tissues. Location/Qualifiers

```

```

source
1. .425
/organism="Mus musculus"
/mol_type="mRNA"
/strain="BALB/c"
/db_xref="taxon:10090"
/clon="G430071G05"
/cell_type="B cells"
/cell_line="CRL-1669 BCL1 Clone 13.20-3B3 cDNA"
BCL1 Clone 13.20-3B3 cDNA"

ORIGIN

Query Match      87.2%; Score 21.8; DB 2; Length 425;
Best Local Similarity 92.0%; Pred. No. 1.2e+02;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAAAAAATCAGCGCCAGGCATGG 25
    |||||
Db 167 AAAAAAAAAATCTTGGCCAGGCATGG 143

RESULT 31
BB789804/c
LOCUS
DEFINITION BB789804 RIKEN full-length enriched, embryo RCB-0549 C1e-H3 cDNA
Mus musculus cDNA clone G430129E01 3', mRNA sequence.
BB789804 425 bp mRNA linear EST 08-JUL-2003
ACCESSION BB789804
VERSION BB789804.1 GI:16958300
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM
Mus musculus

REFERENCE
AUTHORS
Akimura,T., Arakawa,T., Carninci,P., Furuno,M., Hanagaki,T.,
Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Imotani,K.,
Iehii,Y., Ito,M., Kawai,J., Kojima,Y., Konno,H., Kouda,M.,
Matsuyama,T., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R.,
Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sakazume,N.,
Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Suzuki,H., Tagawa,A., Tanaka,T., Tanaka,T., Tanaka,T.,
Tanaka,T., Tomaru,A., Toyota,T., Watahiki,A., Yasunishi,A.,
Muramatsu,M. and Hayashizaki,Y.
RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al.
2001)

TITLE
Unpublished (2001)
JOURNAL
COMMENT
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. . 10 (10), 1617-1630 (2000)
wagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T.,
Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A.
and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
. 10 (11), 1757-1771 (2000)
Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P.,
Sugahara,Y. and Hayashizaki,Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. . 11 (2), 281-289 (2001)
Please visit our web site (http://genome.gsc.riken.go.jp) for
further details.
e mouse tissues. Location/Qualifiers

```

further details.

e mouse tissues.
 Location/Qualifiers
 1. .425
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="G430129E01"
 /cell_line="RCB-0549 Cle-H3"
 /dev_stage="embryo"
 /clone_lib="RIKEN full-length enriched, embryo RCB-0549 Cle-H3 cDNA"
 Query Match 87.2%; Score 21.8; DB 2; Length 425;
 Best Local Similarity 92.0%; Pred. No. 1.2e+02;
 Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 AAAAAAAAAATCAGCGCAGGCATGG 25
 Db 170 AAAAAAAAAATCTGGCCAGGCATGG 146

ORIGIN

Query Match 87.2%; Score 21.8; DB 2; Length 425;
 Best Local Similarity 92.0%; Pred. No. 1.2e+02;
 Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

RESULT 32

BY383430/c
 LOCUS
 DEFINITION
 RIKEN full-length enriched, 4 days embryo whole body Mus musculus cDNA clone I0C0039E02 3', mRNA sequence.

ACCESSION

BY383430.1 GI:26612918

KEYWORDS

EST.

SOURCE

Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridea; Muridae; Murinae; Mus.

REFERENCE

AUTHORS
 Okazaki, I., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nikaide, I., Oatono, N., Saito, R., Suzuki, H., Yamanaka, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C., Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H., Batilov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusci, V., Chothia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A., Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W.J., Pertea, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S., Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M., Sandelin, A., Schneider, C., Semple, C.A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M., Verardo, R., Wagner, L.G., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Inotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S., Rogers, J., Birney, E. and Hayashizaki, Y.
 Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

TITLE

Analysis of the mouse transcriptome based on functional annotation

JOURNAL

Nature 420, 563-573 (2002)

PUBMED

12466851

COMMENT

Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

FEATURES

source

Location/Qualifiers
 1. .426
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="I0C0039E02"
 /tissue_type="whole body"
 /dev_stage="4 days embryo"
 /clone_lib="RIKEN full-length enriched, 4 days embryo whole body"

ORIGIN

Query Match 87.2%; Score 21.8; DB 5; Length 426;
 Best Local Similarity 92.0%; Pred. No. 1.2e+02;
 Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY

1 AAAAAAAAAATCAGCGCAGGCATGG 25

Db

172 AAAAAAAAAATCTGGCCAGGCATGG 148

RESULT 33

AA744879/c

LOCUS

DEFINITION
 NCI-CGAP GCBI Homo sapiens cDNA clone IMAGE:1283781 3' similar to contains Alu repetitive element; contains L1.t2 L1 repetitive element ;, mRNA sequence.

ACCESSION

AA744879.1 GI:2783643

VERSION

EST.

KEYWORDS

Homo sapiens (human)

SOURCE

Homo sapiens

ORGANISM

Homo sapiens

REFERENCE

1 (bases 1 to 431)

AUTHORS

NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

TITLES

National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

JOURNAL

Unpublished (1997)

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: cgaps-r@mail.nih.gov

Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman, Ph.D., Gerald Marti, M.D.

cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/

Alizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S.,

Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H.,

Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R.,

Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K.,

Shiraki, T., Tagami, M., Waki, K., Watahiki, A., Muramatsu, M. and

Hayashizaki, Y. Direct Submission

Computational Analysis of Full-length Mouse cDNAs Compared with

Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new

genes. Genome Res. 10 (10), 1617-1630 (2000)

RIKEN integrated sequence analysis (RISA) system--384-format

sequencing pipeline with 384 multicapillary sequencer. Genome Res.

10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA

encyclopedia: real-time sequence clustering for construction of a

nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

cDNA library was prepared and sequenced in Mouse Genome

Encyclopedia Project of Genome Exploration Research Group in Riken

Genomic Sciences Center and Genome Science Laboratory in RIKEN.

Division of Experimental Animal Research in Riken contributed to

prepare mouse tissues.

Please visit our web site (http://genome.gsc.riken.go.jp) for

further details.

Location/Qualifiers

1. .426

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="I0C0039E02"

/tissue_type="whole body"

/dev_stage="4 days embryo"

/clone_lib="RIKEN full-length enriched, 4 days embryo

whole body"

Bonaldo, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 664 Std Error: 0.00
Seq primer: -40m13 fwd. RT from Amersham
High quality sequence stop: 424.
Location/Qualifiers

FEATURES

source
1. .431
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1283781"
/tissue_type="germinal center B cell"
/lab_host="DH10B"
/clone_lib="NCI CGAP GCB1"
/notes="Vector: p773D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was prepared from human tonsillar cells enriched for germinal center B cells by flow sorting (CD20+, IgD-), provided by Dr. Louis M. Staudt (NCI), Dr. David Allman (NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was primed with a Not I - oligo(dT) primer
[5'-TGTTACCAATCTGAAGTGGAGCGCGCTCATTTTTTTTTTTT-3'
]. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified p773 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 87.2%; Score 21.8; DB 1; Length 431;
Best Local Similarity 92.0%; Pred. No. 1.2e+02;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AAAAAAAAAACAGCCGCGCATGG 25
|||||
Db 431 AAAAAAAAAACCCAGCCGCGCATGG 407

RESULT 34
AI642924 431 bp mRNA linear EST 29-APR-1999
LOCUS v652408.x1 Knowles Solter mouse blastocyst B1 Mus musculus
DEFINITION clone IMAGE:805142 3' similar to gb:L42293 Mus musculus acyl-coenzyme A:cholesterol acyltransferase (MOUSE);, mRNA sequence.

ACCESSION AI642924
VERSION AI642924.1 GI:4721399
KEYWORDS EST.
SOURCE Mus musculus
ORGANISM Mus musculus (house mouse)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Mus.
1 (bases 1 to 431)

REFERENCE
AUTHORS Marra, M., Hallier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.
The WashU-NCI Mouse EST Project 1999

TITLE Unpublished (1999)
JOURNAL Contact: Marra M/WashU-NCI Mouse EST Project 1999
COMMENT Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.

MGI:481486

This clone was previously sequenced on the 5' end only, this new data is from the 3' end
Seq primer: Primer name ambiguous
High quality sequence stop: 422.
Location/Qualifiers

FEATURES

source
1. .431
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J x DBA/2J F1"
/db_xref="taxon:10090"
/clone="IMAGE:805142"
/tissue_type="blastocyst"
/dev_stage="embryo (pre-implantation)"
/lab_host="DH10B"
/clone_lib="Knowles Solter mouse blastocyst B1"
/notes="Organ: embryo; Vector: pSPORT; site 1: NotI; Site 2: Sali; Cloned unidirectionally from mRNA prepared from 800 blastocysts. Primer: Sali (dT): 5'-CGTCCACGCTCGACGCTTTTTTTTTT-3'. cDNAs were cloned into the NotI/Sali sites of a pSPORT vector (Life Technologies). Two different size selections: B1 (larger inserts) and B3."

ORIGIN

Query Match 87.2%; Score 21.8; DB 1; Length 431;
Best Local Similarity 92.0%; Pred. No. 1.2e+02;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AAAAAAAAAATCAGCCGCGCATGG 25
|||||
Db 259 AAAAAAAAAATCTTGGCCGCGCATGG 283

RESULT 35

BY693822/c

LOCUS

DEFINITION

musculus cDNA clone I42029M24 3', mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Mus musculus (house mouse)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

Sciurognathi; Muroidea; Muridae; Murinae; Mus.

1 (bases 1 to 434)

Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,

Nikaido, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I.,

Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,

Schönbach, C., Cojocari, T., Baldarelli, R., Hill, D. P., Bult, C.,

Hume, D. A., Quackenbush, J., Schriml, L. M., Kanapin, A., Matsuda, H.,

Batalov, S., Beisel, K. W., Blake, J. A., Bradt, D., Bruscia, V.,

Chothia, C. F., Cousins, S., Dalla, E., Dragani, T. A.,

Fletcher, C. F., Forrest, A., Frazer, K. S., Gaasterland, T.,

Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S.,

Gustincich, S., Hirokawa, N., Jackson, I. J., Jarvis, E. D., Kanai, A.,

Kawaji, H., Kawasawa, Y., Redziarski, R. M., King, B. L., Konagaya, A.,

Kurochkin, I. V., Lee, Y., Lenhard, B., Lyons, P. A., Maglott, D. R.,

Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T.,

Numata, K., Okido, T., Pavan, W. J., Pertea, G., Pesole, G.,

Petrovsky, N., Pillai, R., Pontius, J. U., Qi, D., Ramachandran, S.,

Ravasi, T., Reed, J. C., Reed, D. J., Reid, J., Ring, B. Z., Ringwald, M.,

Sandelin, A., Schneider, C., Sempile, C. A., Setou, M., Shimada, K.,

Sultana, R., Takenaka, Y., Taylor, M. S., Teasdale, R. D., Tomita, M.,

Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y.,

Wells, C., Wilming, L. G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I.,

Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P.,

Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M.,

Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Alzawa, K.,

Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y.,

Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K.,

Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S., Rogers, J., Birney, E. and Hayashizaki, Y.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
12466851

TITLE

JOURNAL

PUBMED

COMMENT

Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216

Email: genome-res@gsc.riken.jp, URL: <http://genome.gsc.riken.jp/>
Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission
Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Tissues were provided by Takashi Ishikawa (Department of Surgery 2 Yokohama City University 3-9 Fukuura, Kanazawa-ku, Yokohama 236-0004 Japan) whose assistance we gratefully acknowledge.
Please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.

FEATURES

source

```
1. .434
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="I420029M24"
/cell_type="osteoclast-like cell"
/clone_lib="RIKEN full-length enriched, osteoclast-like cell"
```

ORIGIN

```
Query Match : 87.2%; Score 21.8; DB 5; Length 434;
Best Local Similarity 92.0%; Pred. No. 1.2e+02;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

QY 1 AAAAAAAAAATCAGCGCCAGGCATGG 25

Db 177 AAAAAAAAAATCTGGCCAGGCATGG 153

RESULT 36

BY384796/c

LOCUS

DEFINITION BY384796 RIKEN full-length enriched, blastocyst Mus musculus cDNA clone IIC0004N17 3', mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

BY384796 436 bp mRNA linear EST 12-DEC-2002
GI:26614284
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE

AUTHORS

1 (bases 1 to 436)

Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nikaido, I., Otsu, N., Saito, R., Suzuki, H., Yamaguchi, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C., Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H., Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusic, V., Chotha, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A., Flath, C., Fraser, K.S., Gaasterland, T., Gustinich, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A., Kawai, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A., Kurochin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R., Maltas, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Nomura, K., Okido, T., Pavan, W.J., Perte, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramchandran, S., Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M., Sandelin, A., Schneider, C., Semple, C.A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M., Verdard, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavoian, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S., Rogers, J., Birney, E. and Hayashizaki, Y.

Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
12466851

TITLE

JOURNAL

PUBMED

COMMENT

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216

Email: genome-res@gsc.riken.jp, URL: <http://genome.gsc.riken.jp/>
Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission
Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)

Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.

FEATURES

Source

```
1. .436
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IIC0004N17"
```


/dev stage="blastocyst"
/clone_lib="RIKEN full-length enriched, blastocyst"

ORIGIN
Query Match 87.2%; Score 21.8; DB 5; Length 436;
Best Local Similarity 92.0%; Pred. No. 1.2e+02;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 AAAAAAAAAATCGGCCAGGCATGG 25
|||
DB 180 AAAAAAAAAATCGGCCAGGCATGG 156

RESULT 37
BB782484/c
LOCUS BB782484 RIKEN full-length enriched, CRL-1751 WEHI 164 cDNA Mus
DEFINITION musculus cDNA clone G430081B13 3', mRNA sequence.

ACCESSION BB782484.1 GI:16943184

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM

REFERENCE
1 (bases 1 to 447)
Akimura,T., Arakawa,T., Carninci,P., Furuno,M., Hanagaki,T., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Imotani,K., Ishii,Y., Ito,M., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Matsuyama,T., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sakazume,N., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Watahiki,A., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.

TITLE
RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al. 2001)

JOURNAL
Unpublished (2001)
COMMENT
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsr.riken.jp, URL: http://genome.gsc.riken.jp/

Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
wagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)

Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara,Y. and Hayashizaki,Y.
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Please visit our web site (http://genome.gsc.riken.go.jp) for further details.

e mouse tissues.

FEATURES
Location/Qualifiers
1. .447
/organism="Mus musculus"
/mol_type="mRNA"
/strain="BALB/c"
/db_xref="taxon:10090"

ORIGIN

Query Match 87.2%; Score 21.8; DB 2; Length 447;
Best Local Similarity 92.0%; Pred. No. 1.2e+02;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 AAAAAAAAAATCGGCCAGGCATGG 25
|||
DB 193 AAAAAAAAAATCGGCCAGGCATGG 169

RESULT 38
BB819202/c

LOCUS BB819202 RIKEN full-length enriched, lung RCB-0558 LLC cDNA Mus

DEFINITION musculus cDNA clone G730048H03 3', mRNA sequence.

ACCESSION BB819202.1 GI:16991831

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM

REFERENCE

AUTHORS

1 (bases 1 to 470)

Akimura,T., Arakawa,T., Carninci,P., Furuno,M., Hanagaki,T., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Imotani,K., Ishii,Y., Ito,M., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Matsuyama,T., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sakazume,N., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Watahiki,A., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.

TITLE

RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al. 2001)

JOURNAL

Unpublished (2001)

COMMENT

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-res@gsr.riken.jp, URL: http://genome.gsc.riken.jp/

Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.

Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)

wagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.

RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)

Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara,Y. and Hayashizaki,Y.

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

Please visit our web site (http://genome.gsc.riken.go.jp) for further details.

e mouse tissues.

FEATURES

Location/Qualifiers

1. .470

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

/db_xref="taxon:10090"
 /clone="G730048H03"
 /tissue_type="lung"
 /cell_line="RCB-0558 LLC"
 /clone_lib="RIKEN full-length enriched, lung RCB-0558 LLC
 cDNA"

ORIGIN

Query Match 87.2%; Score 21.8; DB 2; Length 470;
 Best Local Similarity 92.0%; Pred. No. 1.2e+02;
 Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAAAAAATCAGCCAGGCATGG 25

Db 218 AAAAAAAAAATCTGGCCAGGCATGG 194

RESULT 39

BM205156 474 bp mRNA linear EST 08-JUN-2003
 LOCUS C0279F09-3 NIA Mouse 7.5-dpc Whole Embryo cDNA Library (Long) Mus
 DEFINITION musculus cDNA clone NIA:C0279F09 IMAGE:30018212 3', mRNA sequence.

ACCESSION BM205156
 VERSION BM205156.2 GI:31529650

KEYWORDS Mus musculus (house mouse)

ORGANISM

SOURCE Mus musculus
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muridea; Muridae; Murinae; Mus.

REFERENCE

PIAO, Y., KO, N.T., LIM, M.K. and KO, M.S.H.
 Construction of long-transcript enriched cDNA libraries from
 submicrogram amounts of total RNAs by a universal PCR amplification
 method

Genome Res. 11 (9), 1553-1558 (2001)

11541199

On Dec 14, 2001 this sequence version replaced gi:17760224.

Other ESTs: C0279F09-5N

Contact: Dawood B. Dudekula

Laboratory of Genetics

National Institute on Aging/National Institutes of Health

333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA

Email: cdna@nigms.nih.gov

Plate: C0279 row: F column: 09

Seq primer: -21M13 Forward

High quality sequence stop: 474

POLYA=Yes.

Location/Qualifiers

1. .474

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="nia:EST:C0279F09-3"

/db_xref="taxon:10090"

/clone="NIA:C0279F09 IMAGE:30018212"

/tissue_type="whole embryo including extraembryonic

tissues at 7.5-days postcoitum"

/dev_stage="7.5-days postcoitum"

/lab_host="DH10B"

/clone_lib="NIA Mouse 7.5-dpc Whole Embryo cDNA Library

(Long)"

/note="Vector: pSPORT1 (Invitrogen); Site 1: SalI; Site 2:

NotI; Mouse cDNA project by the Laboratory of Genetics,

National Institute on Aging (NIA), Intramural Research

Program, NIH (http://nigms.nih.gov/cDNA). This is

a long-transcript enriched cDNA library (Ref. Genome Res.

11: 1553-1558 (2001). [PMID: 11541199]). Total RNAs were

extracted from a pool of four embryos at 7.5-days

postcoitum. Double-stranded cDNAs were synthesized with an

Oligo(dT) primer [Invitrogen].

5'-pGACTAGTCTAGATCGAGGCGCCCTTTT-3' from

7 ug of total RNA, treated with T4 DNA polymerase, and

purified by ethanol-precipitation. The cDNAs were ligated
 to Lone-linker L1-Sal4, purified by phenol/chloroform, and
 separated from free linkers by Centricon 100. Then, the
 cDNAs were amplified by long-range high fidelity PCR using
 Ex Taq polymerase (Takara) with a primer Sal4-S. The
 products were purified by phenol/chloroform and Centricon
 100. The cDNAs were digested with SalI and NotI enzymes
 and cloned into SalI/NotI site of pSPORT1 plasmid vector.
 The DH10B E. coli host was transformed with the ligation
 mixture by the standard chemical method. The average
 insert size is about 2.2 kb. The library was constructed
 by Yulan Piao (NIA)."

ORIGIN

Query Match 87.2%; Score 21.8; DB 3; Length 474;

Best Local Similarity 92.0%; Pred. No. 1.2e+02;

Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAAAAAATCAGCCAGGCATGG 25

Db 253 AAAAAAAAAATCTGGCCAGGCATGG 277

RESULT 40

AZ450659/C

LOCUS

DEFINITION

1M0249D01R Mouse 10kb plasmid UUGC1M library Mus musculus genomic

clone UUGC1M0249D01 R, genomic survey sequence.

ACCESSION AZ450659

VERSION AZ450659.1 GI:10605669

KEYWORDS GSS.

SOURCE Mus musculus (house mouse)

ORGANISM

Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muroidea; Muridae; Murinae; Mus.

REFERENCE

AUTHORS

1 (bases 1 to 491)

Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,

Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T.,

Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von

Niederhausern, A. and Wright, D., Weiss, R.

Mouse whole genome scaffolding with paired end reads from 10kb

plasmid inserts

Unpublished (2000)

Contact: Robert B. Weiss

University of Utah Genome Center

University of Utah

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT

84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00

Plate: 0249 row: D column: 01

Seq primer: CACACAGGAACACGCTATGACC

Class: plasmid ends

High quality sequence stop: 491.

FEATURES

Location/Qualifiers

1. .491

/organism="Mus musculus"

/mol_type="genomic DNA"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="UUGC1M0249D01"

/sex="Male"

/lab_host="E. Coli strain XL10-Gold, TI-resistant, F-"

/clone_lib="Mouse 10kb plasmid UUGC1M library"

/note="Vector: pW42nv; Purified genomic DNA from M.

musculus C57BL/6J (male) was obtained from the Jackson

Laboratory Mouse DNA Resource

(http://www.jax.org/resources/documents/dnares/). The DNA

was hydrodynamically sheared by repeated passage through a

0.005 inch orifice at constant velocity. The sheared DNA

was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

ORIGIN

Query Match 87.2%; Score 21.8; DB 9; Length 491;
Best Local Similarity 92.0%; Pred. No. 1.2e+02;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAAAAAATCAGCGCAGCATGG 25
|||||
Db 120 AAAAAAAAAATCTGGCCAGGCATGG 96

RESULT 41

BM203574
LOCUS
DEFINITION
C0256E04-3 NIA Mouse 7.5-dpc Whole Embryo cDNA Library (Long) Mus musculus cDNA clone NIA:C0256E04 IMAGE:30015987 3', mRNA sequence.

ACCESSION

BM203574

VERSION

BM203574.2

GI:31528419

EST.

Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 514)

Piao,Y., Ko,N.T., Lim,M.K. and Ko,M.S.H.

Construction of long-transcript enriched cDNA libraries from submicrogram amounts of total RNAs by a universal PCR amplification method

Genome Res. 11 (9), 1553-1558 (2001)

1154199

On Dec 14, 2001 this sequence version replaced gi:17758197.

Other ESTs: C0256E04-5N

Contact: Dawood B. Dudekula

Laboratory of Genetics

National Institute on Aging/National Institutes of Health

333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA

Email: cdna@lgsun.grc.nia.nih.gov

Plate: C0256 row: E column: 04

Seq primer: -21M13 Forward

High quality sequence stop: 514

POLYA=Yes.

Location/Qualifiers

1. 514

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="niaEST:C0256E04-3"

/db_xref="taxon:10090"

/clone="NIA:C0256E04 IMAGE:30015987"

/tissue_type="whole embryo including extraembryonic tissues at 7.5-days postcoitum"

/dev_stage="7.5-days postcoitum"

/lab_host="DH10B"

/clone_lib="NIA Mouse 7.5-dpc Whole Embryo cDNA Library (Long)"

/note="Vector: pSPORT1 (Invitrogen); Site 1: SalI; Site 2: NotI; Mouse cDNA project by the Laboratory of Genetics, National Institute on Aging (NIA), Intramural Research Program, NIH (<http://lgsun.grc.nia.nih.gov/cdna>). This is

a long-transcript enriched cDNA library (Ref. Genome Res. 11: 1553-1558 (2001). [PMID: 11544199]). Total RNAs were extracted from a pool of four embryos at 7.5-days postcoitum. Double-stranded cDNAs were synthesized with an Oligo(dT) primer [Invitrogen: 5'-pactagttctagatcgacgagcgccctttttttttt-3'] from 7 ug of total RNA, treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to Lone-linker LL-Sal4, purified by phenol/chloroform, and separated from free linkers by Centricon 100. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer Sal4-S. The products were purified by phenol/chloroform and Centricon 100. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pSPORT1 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by the standard chemical method. The average insert size is about 2.2 kb. The library was constructed by Yulan Piao (NIA)."

ORIGIN

Query Match 87.2%; Score 21.8; DB 3; Length 514;
Best Local Similarity 92.0%; Pred. No. 1.2e+02;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAAAAAATCAGCGCAGGCATGG 25
|||||
Db 252 AAAAAAAAAATCTGGCCAGGCATGG 276

RESULT 42

AA721682/c

LOCUS

DEFINITION

AA721682

519 bp mRNA linear EST 18-FEB-1998

AA721682

GI:2736592

EST.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 519)

NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: cgaps-r@mail.nih.gov

Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman, Ph.D., Gerald Marti, M.D.

cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.

cDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/IMG at:

www-bio.llnl.gov/bbrp/image/image.html

Insert Length: 894 Std Error: 0.00

Seq primer: -40m13 fwd. ET from Amersham

High quality sequence stop: 440.

Location/Qualifiers

1. 519

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:1283873"

/tissue_type="germinal center B cell"

/lab_host="DH10B"

/clone_lib="NCI_CGAP_GCB1"

/note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was prepared from human tonsillar cells enriched for germinal center B cells by flow sorting (CD20+, Igd-), provided by Dr. Louis M. Staudt (NCI), Dr. David Allman (NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was primed with a Not I - oligo(dT) primer
[5'-TGTTACCAATCGAGTGGAGCGCGCCCTCATTTTTTTTTTTT-3',
]. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 87.2%; Score 21.8; DB 1; Length 519;
Best Local Similarity 92.0%; Pred. No. 1.2e+02;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAAAAAATCAGCGCCAGGCATGG 25

DB 430 AAAAAAAAAATCAGCGCCAGGCATGG 406

RESULT 43

BB753963/c

LOCUS

DEFINITION

BB753963 RIKEN full-length enriched, melanocyte Mus musculus cDNA

clone G270038E14 3', mRNA sequence.

ACCESSION

BB753963

VERSION

BB753963.1

KEYWORDS

EST.

SOURCE

Mus musculus (house mouse)

ORGANISM

Mus musculus

REFERENCE

AUTHORS

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

Sciurognathi; Murioidea; Muridae; Murinae; Mus.

1 (bases 1 to 555)

Akimura,T., Arakawa,T., Carninci,P., Furuno,M., Hanagaki,T.,

Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Imotani,K.,

Ishii,Y., Ito,M., Kawai,J., Kojima,Y., Konno,H., Kouda,M.,

Matsumura,T., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R.,

Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sakazume,N.,

Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,

Sogabe,Y., Suzuki,H., Tagawa,A., Takanashi,F., Takaku-Akahira,S.,

Tanaka,T., Tomaru,A., Toya,T., Watahiki,A., Yasunishi,A.,

Muramatsu,M. and Hayashizaki,Y.

RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al.

2001)

Unpublished (2001)

Contact: Yoshinide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic

Sciences Center (GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-res@gs.riken.jp, URL: http://genome.gsc.riken.jp/

Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,

Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new

genes. Genome Res. 10 (10), 1617-1630 (2000)

wagi,K., Fujiwaki,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,

Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T.,

Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A.

and Hayashizaki,Y.

RIKEN integrated sequence analysis (RISA) system--384-format

sequencing pipeline with 384 multicapillary sequencer. Genome Res.

encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Please visit our web site (http://genome.gsc.riken.go.jp) for
further details.

e mouse

tissues.

Location/Qualifiers

source

1. 555

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="G270038E14"

/cell_type="melanocyte"

/clone_lib="RIKEN full-length enriched, melanocyte"

ORIGIN

Query Match 87.2%; Score 21.8; DB 2; Length 555;
Best Local Similarity 92.0%; Pred. No. 1.2e+02;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAAAAAATCAGCGCCAGGCATGG 25

DB 402 AAAAAAAAAATCTTGGCCAGGCATGG 378

RESULT 44

BP768879

LOCUS

DEFINITION

BP768879 mouse (C57BL/6) pancreatic islet library with

recombination-based method Mus musculus cDNA clone mid39036 3',

mRNA sequence.

ACCESSION

BP768879

VERSION

BP768879.1

KEYWORDS

EST.

SOURCE

Mus musculus (house mouse)

ORGANISM

Mus musculus

REFERENCE

AUTHORS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

Sciurognathi; Murioidea; Muridae; Murinae; Mus.

1 (bases 1 to 575)

Nishimura,M., Yokoi,N., Miki,T., Horikawa,Y., Yoshioka,H.,

Takeda,J., Ohara,O. and Seino,S.

Construction of a multi-functional cDNA library specific for mouse

pancreatic islets and its application to microarray

DNA Res. 11 (5), 315-323 (2004)

PUBMED

15747579

COMMENT

Contact: Susumu Seino

Division of Cellular and Molecular Medicine

Kobe University Graduate School of Medicine

7-5-1 Kusunoki-cho, Chuo-ku, Kobe, Hyogo 650-0017, Japan

Tel: 81-78-382-5360

Fax: 81-78-382-5370

Email: seino@med.kobe-u.ac.jp.

Location/Qualifiers

source

1. 575

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6"

/db_xref="taxon:10090"

/clone="mid39036"

/sex="male"

/tissue_type="pancreatic islet"

/dev_stage="adult"

/clone_lib="mouse (C57BL/6) pancreatic islet library with

recombination-based method"

ORIGIN

Query Match 87.2%; Score 21.8; DB 3; Length 575;
Best Local Similarity 92.0%; Pred. No. 1.2e+02;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAAAAAATCAGCGCCAGGCATGG 25

|||||

|||||

|||||

|||||

|||||

|||||

```
Db 312 AAAAAAAAAATCTTGGCCAGGCATGG 336

RESULT 45
CK345513 593 bp mRNA linear EST 22-DEC-2003
LOCUS L0247D01-3 NIA Mouse Newborn Ovary cDNA Library Mus musculus cDNA
DEFINITION clone L0247D01 3', mRNA sequence.
ACCESSION CK345513
VERSION CK345513.1 GI:40301126
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
TITLE Sciurognathi; Muroidea; Muridae; Murinae; Mus.
1 (bases 1 to 593)
Piao,Y., Ko,N.T., Lim,M.K. and Ko,M.S.H.
Construction of long-transcript enriched cDNA libraries from
submicrogram amounts of total RNAs by a universal PCR amplification
method
JOURNAL Genome Res. 11 (9), 1553-1558 (2001)
PUBMED 11544199
COMMENT Contact: Dawood B. Dudekula
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@lgsun.grc.nia.nih.gov
Plate: L0247 row: D column: 01
Seq primer: -21M13 Forward
High quality sequence stop: 593
POLYA=Yes.

FEATURES
source Location/Qualifiers
1..593
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="niaEST:L0247D01-3"
/db_xref="taxon:10090"
/clone="L0247D01"
/sex="female"
/dev_stage="Newborn Ovary"
/lab_host="DH10B"
/clone_lib="NIA Mouse Newborn Ovary cDNA Library"
/notes="Vector: pSPORT1 (Gibco/BRL Life Technology);
Site 1: SalI; Site 2: NotI; Total RNAs were extracted from
7 Newborn Ovary. The double-stranded cDNA was synthesized
by Gibco's kit with an Oligo(dT) primer [NotI
primer-adaptor from GibcoBRL]
[5'-pGACTAGTTCTAGATCGGCGCGCCCTTTT-3'] from
2.56ug of total RNA. The double-stranded cDNAs were
treated with T4 DNA polymerase and purified by
ethanol-precipitation. The cDNAs were ligated to
Lone-linker L1-Sal3 (include SalI sequence). The cDNAs
were purified by phenol/chloroform and separated from
free linkers by Centricon 100. Then, cDNAs were amplified
by long-range high fidelity PCR using Takara's Ex Taq
polymerase. Then, the cDNAs were purified by
phenol/chloroform and by Centricon 100. The cDNAs were
digested with SalI and NotI enzymes. Then, the cDNAs were
size selected by Gibco's Size Fractionation Column. The
cDNAs were cloned into SalI/NotI site of pSPORT1 plasmid
vector. The DH10B E. coli host was transformed with the
ligation mixture by chemical method. The library was
constructed by Xiaohong Wang and Yulan Piao."

ORIGIN
Query Match 87.2%; Score 21.8; DB 7; Length 593;
Best Local Similarity 92.0%; Pred. No. 1.2e+02;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AAAAAAAAAATCTTGGCCAGGCATGG 25
||||||| |||||||
```

```
Db 254 AAAAAAAAAATCTTGGCCAGGCATGG 278

RESULT 46
AW554397 595 bp mRNA linear EST 31-AUG-2000
LOCUS L0240B05-3 NIA Mouse Newborn Ovary cDNA Library Mus musculus cDNA
DEFINITION clone L0240B05 3', mRNA sequence.
ACCESSION AW554397
VERSION AW554397.1 GI:7199820
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
TITLE Sciurognathi; Muroidea; Muridae; Murinae; Mus.
1 (bases 1 to 595)
Tanaka,T.S., Jaradat,S.A., Lim,M.K., Kargul,G.J., Wang,X.,
Grahovac,M.J., Pantano,S., Sano,Y., Piao,Y., Nagaraja,R., Doi,H.,
Wood,W.H. III, Becker,K.G. and Ko,M.S.H.
Genome-wide expression profiling of mid-gestation placenta and
embryo using a 15,000 mouse developmental cDNA microarray
Proc. Natl. Acad. Sci. U.S.A. 97 (16), 9127-9132 (2000)
JOURNAL 10922068
PUBMED
COMMENT Contact: George J. Kargul
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@lgsun.grc.nia.nih.gov
Plate: L0240 row: B column: 05
Seq primer: -21M13 Forward
High quality sequence stop: 595
POLYA=Yes.

FEATURES
source Location/Qualifiers
1..595
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="niaEST:L0240B05-3"
/db_xref="taxon:10090"
/clone="L0240B05"
/sex="female"
/dev_stage="Newborn Ovary"
/lab_host="DH10B"
/clone_lib="NIA Mouse Newborn Ovary cDNA Library"
/notes="Vector: pSPORT1 (Gibco/BRL Life Technology);
Site 1: SalI; Site 2: NotI; Total RNAs were extracted from
7 Newborn Ovary. The double-stranded cDNA was synthesized
by Gibco's kit with an Oligo(dT) primer [NotI
primer-adaptor from GibcoBRL]
[5'-pGACTAGTTCTAGATCGGCGCGCCCTTTT-3'] from
2.56ug of total RNA. The double-stranded cDNAs were
treated with T4 DNA polymerase and purified by
ethanol-precipitation. The cDNAs were ligated to
Lone-linker L1-Sal3 (include SalI sequence). The cDNAs
were purified by phenol/chloroform and separated from
free linkers by Centricon 100. Then, cDNAs were amplified
by long-range high fidelity PCR using Takara's Ex Taq
polymerase. Then, the cDNAs were purified by
phenol/chloroform and by Centricon 100. The cDNAs were
digested with SalI and NotI enzymes. Then, the cDNAs were
size selected by Gibco's Size Fractionation Column. The
cDNAs were cloned into SalI/NotI site of pSPORT1 plasmid
vector. The DH10B E. coli host was transformed with the
ligation mixture by chemical method. The library was
constructed by Xiaohong Wang and Yulan Piao."

ORIGIN
Query Match 87.2%; Score 21.8; DB 1; Length 595;
Best Local Similarity 92.0%; Pred. No. 1.2e+02;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AAAAAAAAAATCTTGGCCAGGCATGG 25
```


Search completed: February 3, 2006, 22:01:35
Job time : 2963.67 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 3, 2006, 11:35:31 ; Search time 80.6667 Seconds
(without alignments)
550.897 Million cell updates/sec

Title: US-10-719-900-1

Perfect score: 25

Sequence: 1 aaaaaaaatcacgcagcgatgg 25

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database : Issued Patents NA:*

- 1: /cgn2_6/ptodata/1/ina/1 COMB.seq:*
- 2: /cgn2_6/ptodata/1/ina/5 COMB.seq:*
- 3: /cgn2_6/ptodata/1/ina/6A COMB.seq:*
- 4: /cgn2_6/ptodata/1/ina/6B COMB.seq:*
- 5: /cgn2_6/ptodata/1/ina/H COMB.seq:*
- 6: /cgn2_6/ptodata/1/ina/PCPUS COMB.seq:*
- 7: /cgn2_6/ptodata/1/ina/PP COMB.seq:*
- 8: /cgn2_6/ptodata/1/ina/RE COMB.seq:*
- 9: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	23.4	93.6	51905	3	US-09-949-002-667
C 2	23.4	93.6	51905	3	US-09-949-002-781
C 3	23.4	93.6	107827	3	US-09-949-016-15790
C 4	21.8	87.2	23640	3	US-09-949-016-13646
C 5	21.8	87.2	27592	3	US-09-949-016-14682
C 6	21.8	87.2	27677	3	US-09-949-016-13055
C 7	21.8	87.2	84558	3	US-09-949-016-17572
C 8	21.8	87.2	174170	3	US-09-949-016-14810
C 9	21.8	87.2	174318	3	US-09-949-016-14811
C 10	21.8	87.2	174318	3	US-09-949-016-11880
C 11	21.8	87.2	174318	3	US-09-949-016-14812
C 12	21.8	87.2	174318	3	US-09-949-016-14813
C 13	20.8	83.2	24139	3	US-09-949-016-14675
C 14	20.8	83.2	125188	3	US-09-949-016-11980
C 15	20.8	80.8	138	3	US-09-513-999C-21508
C 16	20.2	80.8	206	3	US-09-513-999C-18230
C 17	20.2	80.8	601	3	US-09-949-016-17632
C 18	20.2	80.8	601	3	US-09-949-016-17633
C 19	20.2	80.8	601	3	US-09-949-016-16195
C 20	20.2	80.8	601	3	US-09-949-016-36196
C 21	20.2	80.8	601	3	US-09-949-016-36197
C 22	20.2	80.8	601	3	US-09-949-016-41289
C 23	20.2	80.8	601	3	US-09-949-016-41290
C 24	20.2	80.8	601	3	US-09-949-016-41290

Sequence 48718, A	US-09-949-016-48718	601	80.8	20.2
Sequence 54603, A	US-09-949-016-54603	601	80.8	20.2
Sequence 54604, A	US-09-949-016-54604	601	80.8	20.2
Sequence 54605, A	US-09-949-016-54605	601	80.8	20.2
Sequence 54635, A	US-09-949-016-54635	601	80.8	20.2
Sequence 54636, A	US-09-949-016-54636	601	80.8	20.2
Sequence 54637, A	US-09-949-016-54637	601	80.8	20.2
Sequence 54667, A	US-09-949-016-54667	601	80.8	20.2
Sequence 54668, A	US-09-949-016-54668	601	80.8	20.2
Sequence 54669, A	US-09-949-016-54669	601	80.8	20.2
Sequence 56301, A	US-09-949-016-56301	601	80.8	20.2
Sequence 56302, A	US-09-949-016-56302	601	80.8	20.2
Sequence 59272, A	US-09-949-016-59272	601	80.8	20.2
Sequence 63344, A	US-09-949-016-63344	601	80.8	20.2
Sequence 119032, A	US-09-949-016-119032	601	80.8	20.2
Sequence 119033, A	US-09-949-016-119033	601	80.8	20.2
Sequence 119034, A	US-09-949-016-119034	601	80.8	20.2
Sequence 119088, A	US-09-949-016-119088	601	80.8	20.2
Sequence 119089, A	US-09-949-016-119089	601	80.8	20.2
Sequence 119070, A	US-09-949-016-119070	601	80.8	20.2
Sequence 119104, A	US-09-949-016-119104	601	80.8	20.2
Sequence 119105, A	US-09-949-016-119105	601	80.8	20.2
Sequence 119176, A	US-09-949-016-119176	601	80.8	20.2
Sequence 119177, A	US-09-949-016-119177	601	80.8	20.2
Sequence 119178, A	US-09-949-016-119178	601	80.8	20.2
Sequence 119212, A	US-09-949-016-119212	601	80.8	20.2
Sequence 119213, A	US-09-949-016-119213	601	80.8	20.2
Sequence 119214, A	US-09-949-016-119214	601	80.8	20.2
Sequence 135443, A	US-09-949-016-135443	601	80.8	20.2
Sequence 135444, A	US-09-949-016-135444	601	80.8	20.2
Sequence 175158, A	US-09-949-016-175158	601	80.8	20.2
Sequence 183999, A	US-09-949-016-183999	601	80.8	20.2
Sequence 184000, A	US-09-949-016-184000	601	80.8	20.2
Sequence 184104, A	US-09-949-016-184104	601	80.8	20.2
Sequence 184105, A	US-09-949-016-184105	601	80.8	20.2
Sequence 184209, A	US-09-949-016-184209	601	80.8	20.2
Sequence 184210, A	US-09-949-016-184210	601	80.8	20.2
Sequence 184314, A	US-09-949-016-184314	601	80.8	20.2
Sequence 184315, A	US-09-949-016-184315	601	80.8	20.2
Sequence 130842, A	US-09-949-016-130842	601	80.8	20.2
Sequence 1726, Ap	US-09-949-002-1726	601	80.8	20.2
Sequence 3436, Ap	US-09-949-002-3436	601	80.8	20.2
Sequence 3437, Ap	US-09-949-002-3437	601	80.8	20.2
Sequence 3801, Ap	US-09-949-002-3801	601	80.8	20.2
Sequence 6416, Ap	US-09-949-002-6416	601	80.8	20.2
Sequence 7692, Ap	US-09-949-002-7692	601	80.8	20.2
Sequence 9453, Ap	US-09-949-002-9453	601	80.8	20.2
Sequence 9454, Ap	US-09-949-002-9454	601	80.8	20.2
Sequence 43, Appli	US-09-511-6258-43	601	80.8	20.2
Sequence 4, Appli	US-09-103-875-4	1110	80.8	20.2
Sequence 14050, A	US-09-949-016-14050	4460	80.8	20.2
Sequence 15564, A	US-09-949-016-15564	5819	80.8	20.2
Sequence 16668, A	US-09-949-016-16668	7482	80.8	20.2
Sequence 12913, A	US-09-949-016-12913	10134	80.8	20.2
Sequence 14446, A	US-09-949-016-14446	10784	80.8	20.2
Sequence 11937, A	US-09-949-016-11937	12921	80.8	20.2
Sequence 16709, A	US-09-949-016-16709	15376	80.8	20.2
Sequence 13708, A	US-09-949-016-13708	15722	80.8	20.2
Sequence 16229, A	US-09-949-016-16229	16621	80.8	20.2
Sequence 5, Appli	US-09-511-6258-5	17013	80.8	20.2
Sequence 666, App	US-09-949-002-666	17425	80.8	20.2
Sequence 17198, A	US-09-949-016-17198	20407	80.8	20.2
Sequence 726, App	US-09-949-002-726	20495	80.8	20.2
Sequence 13500, A	US-09-949-016-13500	23015	80.8	20.2
Sequence 16374, A	US-09-949-016-16374	23953	80.8	20.2
Sequence 17542, A	US-09-949-016-17542	24154	80.8	20.2
Sequence 15191, A	US-09-949-016-15191	24167	80.8	20.2
Sequence 15192, A	US-09-949-016-15192	27968	80.8	20.2
Sequence 14836, A	US-09-949-016-14836	27968	80.8	20.2
Sequence 14836, A	US-09-949-016-14836	29994	80.8	20.2

```
c 98 20.2 80.8 31166 3 US-09-949-016-13464 Sequence 13464, A
c 99 20.2 80.8 34539 3 US-09-949-016-12226 Sequence 12226, A
c 100 20.2 80.8 34540 3 US-09-949-016-13156 Sequence 13156, A
c 101 20.2 80.8 36240 3 US-09-949-016-13176 Sequence 13176, A
c 102 20.2 80.8 36907 3 US-09-949-002-749 Sequence 749, App
c 103 20.2 80.8 37875 3 US-09-949-016-13182 Sequence 13182, A
c 104 20.2 80.8 43690 3 US-09-949-016-13904 Sequence 13904, A
c 105 20.2 80.8 44342 3 US-09-949-016-12661 Sequence 12661, A
c 106 20.2 80.8 44377 3 US-09-949-016-11840 Sequence 11840, A
c 107 20.2 80.8 44378 3 US-09-949-016-15969 Sequence 15969, A
c 108 20.2 80.8 45546 3 US-09-146-053-6 Sequence 6, Appl
c 109 20.2 80.8 47493 3 US-09-949-016-13241 Sequence 13241, A
c 110 20.2 80.8 50836 3 US-09-949-016-16722 Sequence 16722, A
c 111 20.2 80.8 51539 3 US-09-949-002-677 Sequence 677, App
c 112 20.2 80.8 51540 3 US-09-949-002-819 Sequence 819, App
c 113 20.2 80.8 55068 3 US-09-949-002-691 Sequence 691, App
c 114 20.2 80.8 55068 3 US-09-949-002-778 Sequence 778, App
c 115 20.2 80.8 58133 3 US-09-949-016-16464 Sequence 16464, A
c 116 20.2 80.8 60194 3 US-09-949-002-656 Sequence 656, App
c 117 20.2 80.8 60195 3 US-09-949-002-697 Sequence 697, App
c 118 20.2 80.8 63783 3 US-09-949-016-13576 Sequence 13576, A
c 119 20.2 80.8 68444 3 US-09-949-016-13968 Sequence 13968, A
c 120 20.2 80.8 75212 3 US-09-949-016-13313 Sequence 13313, A
c 121 20.2 80.8 75212 3 US-09-949-016-13314 Sequence 13314, A
c 122 20.2 80.8 75212 3 US-09-949-016-13315 Sequence 13315, A
c 123 20.2 80.8 78491 3 US-09-949-016-15132 Sequence 15132, A
c 124 20.2 80.8 78720 3 US-09-949-016-12710 Sequence 12710, A
c 125 20.2 80.8 78720 3 US-09-949-016-17283 Sequence 17283, A
c 126 20.2 80.8 82000 3 US-09-949-016-15595 Sequence 15595, A
c 127 20.2 80.8 84558 3 US-09-949-016-15752 Sequence 15752, A
c 128 20.2 80.8 87352 3 US-09-949-016-12053 Sequence 12053, A
c 129 20.2 80.8 87352 3 US-09-949-016-12721 Sequence 12721, A
c 130 20.2 80.8 87352 3 US-09-949-016-15692 Sequence 15692, A
c 131 20.2 80.8 87352 3 US-09-949-016-15693 Sequence 15693, A
c 132 20.2 80.8 87470 3 US-09-949-016-15881 Sequence 15881, A
c 133 20.2 80.8 87523 3 US-09-949-016-12670 Sequence 12670, A
c 134 20.2 80.8 87523 3 US-09-949-016-15047 Sequence 15047, A
c 135 20.2 80.8 87523 3 US-09-949-016-15048 Sequence 15048, A
c 136 20.2 80.8 87523 3 US-09-949-016-15049 Sequence 15049, A
c 137 20.2 80.8 87869 3 US-09-949-016-11744 Sequence 11744, A
c 138 20.2 80.8 87869 3 US-09-949-016-15044 Sequence 15044, A
c 139 20.2 80.8 87869 3 US-09-949-016-15045 Sequence 15045, A
c 140 20.2 80.8 87869 3 US-09-949-016-15046 Sequence 15046, A
c 141 20.2 80.8 106746 3 US-09-326-402C-1 Sequence 1, Appl
c 142 20.2 80.8 106746 3 US-09-326-402C-12 Sequence 12, Appl
c 143 20.2 80.8 107085 3 US-09-949-016-13157 Sequence 13157, A
c 144 20.2 80.8 131332 3 US-09-949-016-15535 Sequence 15535, A
c 145 20.2 80.8 132871 3 US-09-949-016-13863 Sequence 13863, A
c 146 20.2 80.8 134292 3 US-09-949-016-12158 Sequence 12158, A
c 147 20.2 80.8 150597 3 US-09-949-016-15379 Sequence 15379, A
c 148 20.2 80.8 161607 3 US-09-949-016-12210 Sequence 12210, A
c 149 20.2 80.8 173992 3 US-09-949-016-13379 Sequence 13379, A
c 150 20.2 80.8 246230 3 US-09-949-016-17019 Sequence 17019, A
```

ALIGNMENTS

```
RESULT 1
US-09-949-002-667/c
; Software: FastSeq for Windows Version 4.0
; SEQ ID NO 667
; LENGTH: 51905
; TYPE: DNA
; ORGANISM: Human
US-09-949-002-667

Query Match 93.6%; Score 23.4; DB 3; Length 51905;
Best Local Similarity 96.0%; Pred. No. 1.2;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAAAAAAAATCAGCGCCAGGCATGG 25
|||||
Db 16760 AAAAAAAAAATCAGCGCCAGGCATGG 16736

RESULT 2
US-09-949-002-781/c
; Software: FastSeq for Windows Version 4.0
; SEQ ID NO 781
; LENGTH: 51905
; TYPE: DNA
; ORGANISM: Human
US-09-949-002-781

Query Match 93.6%; Score 23.4; DB 3; Length 51905;
Best Local Similarity 96.0%; Pred. No. 1.2;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAAAAAAAATCAGCGCCAGGCATGG 25
|||||
Db 16760 AAAAAAAAAATCAGCGCCAGGCATGG 16736

RESULT 3
US-09-949-016-15790
; Software: FastSeq for Windows Version 4.0
; SEQ ID NO 15790
; LENGTH: 107827
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15790

Query Match 93.6%; Score 23.4; DB 3; Length 107827;
Best Local Similarity 96.0%; Pred. No. 1.2;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAAAAAAAATCAGCGCCAGGCATGG 25
|||||
Db 16760 AAAAAAAAAATCAGCGCCAGGCATGG 16736

RESULT 3
US-09-949-016-15790
; Software: FastSeq for Windows Version 4.0
; SEQ ID NO 15790
; LENGTH: 107827
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15790

Query Match 93.6%; Score 23.4; DB 3; Length 107827;
Best Local Similarity 96.0%; Pred. No. 1.2;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAAAAAAAATCAGCGCCAGGCATGG 25
|||||
Db 16760 AAAAAAAAAATCAGCGCCAGGCATGG 16736

GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL000790
; CURRENT APPLICATION NUMBER: US/09/949,002
; CURRENT FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/231,401
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 10823
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 781
; LENGTH: 51905
; TYPE: DNA
; ORGANISM: Human
US-09-949-002-781

GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL000790
; CURRENT APPLICATION NUMBER: US/09/949,002
; CURRENT FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/231,401
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 10823
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15790
; LENGTH: 107827
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15790

GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15790
; LENGTH: 107827
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15790

GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL000790
; CURRENT APPLICATION NUMBER: US/09/949,002
; CURRENT FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/231,401
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 10823
```

```
Best Local Similarity 96.0%; Pred. No. 1.3;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AAAAAAAAAATCAGCGCCAGGCATGG 25
    ||||| ||||| ||||| ||||| |||||
Db 69754 AAAAAAAAAATCAGCGCCAGGCATGG 69778

RESULT 4
US-09-949-016-13646
; Sequence 13646, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13646
; LENGTH: 23640
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(23640)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13646

Query Match 87.2%; Score 21.8; DB 3; Length 23640;
Best Local Similarity 92.0%; Pred. No. 5.3;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AAAAAAAAAATCAGCGCCAGGCATGG 25
    ||||| ||||| ||||| ||||| |||||
Db 18856 AAAAAAAAAATCAGCGCCAGGCATGG 18880

RESULT 5
US-09-949-016-14682
; Sequence 14682, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14682
; LENGTH: 27592
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(27592)
; OTHER INFORMATION: n = A,T,C or G

Best Local Similarity 96.0%; Pred. No. 1.3;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AAAAAAAAAATCAGCGCCAGGCATGG 25
    ||||| ||||| ||||| ||||| |||||
Db 69754 AAAAAAAAAATCAGCGCCAGGCATGG 69778

RESULT 6
US-09-949-016-13055
; Sequence 13055, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13055
; LENGTH: 27767
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(27767)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13055

Query Match 87.2%; Score 21.8; DB 3; Length 27767;
Best Local Similarity 92.0%; Pred. No. 5.4;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AAAAAAAAAATCAGCGCCAGGCATGG 25
    ||||| ||||| ||||| ||||| |||||
Db 26968 AAAAAAAAAATCAGCGCCAGGCATGG 26992

RESULT 7
US-09-949-016-15752/c
; Sequence 15752, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15752
; LENGTH: 84558
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(84558)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15752
```

```
Query Match      87.2%; Score 21.8; DB 3; Length 84558;
Best Local Similarity 92.0%; Pred. No. 6.2;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAAAAAATCAGGCCAGGCATGG 25
   ||||| ||||| ||||| ||||| |||||
Db 14022 AAAAAAAAAACAGCCAGGCATGG 13998

RESULT 8
US-09-949-016-14810
; Sequence 14810, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14810
; LENGTH: 174170
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(174170)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14810

Query Match      87.2%; Score 21.8; DB 3; Length 174170;
Best Local Similarity 92.0%; Pred. No. 6.8;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAAAAAATCAGGCCAGGCATGG 25
   ||||| ||||| ||||| ||||| |||||
Db 14074 AAAAAAAAAATCAGGCCAGGCATGG 14098

RESULT 10
US-09-949-016-11880
; Sequence 11880, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11880
; LENGTH: 174318
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(174318)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-11880

Query Match      87.2%; Score 21.8; DB 3; Length 174318;
Best Local Similarity 92.0%; Pred. No. 6.8;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAAAAAATCAGGCCAGGCATGG 25
   ||||| ||||| ||||| ||||| |||||
Db 14222 AAAAAAAAAATCAGGCCAGGCATGG 14246

RESULT 11
US-09-949-016-14812
; Sequence 14812, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14812
; LENGTH: 174318
; TYPE: DNA
```

```
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(174318)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14812

Query Match      87.2%; Score 21.8; DB 3; Length 174318;
Best Local Similarity 92.0%; Pred. No. 6.8;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AAAAAAAAAATCAGGCCAGGCATGG 25
   |||||
Db 14222 AAAAAAAAAATCAGGCCAGGCATGG 14246

RESULT 12
US-09-949-016-14813
; Sequence 14813, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14813
; LENGTH: 174318
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(174318)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14813

Query Match      87.2%; Score 21.8; DB 3; Length 174318;
Best Local Similarity 92.0%; Pred. No. 6.8;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AAAAAAAAAATCAGGCCAGGCATGG 25
   |||||
Db 14222 AAAAAAAAAATCAGGCCAGGCATGG 14246

RESULT 13
US-09-949-016-14675
; Sequence 14675, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14675
; LENGTH: 24139
; TYPE: DNA
; ORGANISM: Human
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14675

Query Match      83.2%; Score 20.8; DB 3; Length 24139;
Best Local Similarity 91.7%; Pred. No. 15;
Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AAAAAAAAAATCAGGCCAGGCATG 24
   |||||
Db 14130 AAAAAAAAAATCAGGCCAGGCATG 14153

RESULT 14
US-09-949-016-11980
; Sequence 11980, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11980
; LENGTH: 125188
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(125188)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-11980

Query Match      83.2%; Score 20.8; DB 3; Length 125188;
Best Local Similarity 91.7%; Pred. No. 18;
Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AAAAAAAAAATCAGGCCAGGCATG 24
   |||||
Db 93166 AAAAAAAAAATCAGGCCAGGCATG 93189

RESULT 15
US-09-513-999C-21508/c
; Sequence 21508, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 21508
; LENGTH: 138
; TYPE: DNA
```


; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17633
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17633

Query Match 80.8%; Score 20.2; DB 3; Length 601;
Best Local Similarity 88.0%; Pred. No. 17;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AAAAAAAAAATCAGCGCCAGGCATGG 25
Db 68 AAAAAAAAAAAAAAAAAAGGCCAGGCATGG 44

RESULT 20

US-09-949-016-36195/c
; Sequence 36195, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 36195
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-36195

Query Match 80.8%; Score 20.2; DB 3; Length 601;
Best Local Similarity 88.0%; Pred. No. 17;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AAAAAAAAAATCAGCGCCAGGCATGG 25
Db 360 AAAAAAAAAAAAAAAAAAGGCCAGGCATGG 336

RESULT 21

US-09-949-016-36196/c
; Sequence 36196, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 36196
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-36196

Query Match 80.8%; Score 20.2; DB 3; Length 601;
Best Local Similarity 88.0%; Pred. No. 17;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AAAAAAAAAATCAGCGCCAGGCATGG 25
Db 137 AAAAAAAAAAAAAAAAAAGGCCAGGCATGG 113

RESULT 22

US-09-949-016-36197/c
; Sequence 36197, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 36197
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-36197

Query Match 80.8%; Score 20.2; DB 3; Length 601;
Best Local Similarity 88.0%; Pred. No. 17;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AAAAAAAAAATCAGCGCCAGGCATGG 25
Db 68 AAAAAAAAAAAAAAAAAAGGCCAGGCATGG 44

RESULT 23

US-09-949-016-41289/c
; Sequence 41289, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 41289
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-41289

```
Query Match      80.8%; Score 20.2; DB 3; Length 601;
Best Local Similarity 88.0%;
Pred. No. 17;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

Qy 1 AAAAAAAAAATCAGGCCAGGCATGG 25
|||
333 AAAAAAAAAAATCGGCCAGGCATGG 309

RESULT 24

US-09-949-016-41290/c

Sequence 41290, Application US/09949016

Patent No. 6812339

GENERAL INFORMATION:

APPLICANT: VENTER, J. Craig et al.

TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

FILE REFERENCE: C0001307

CURRENT APPLICATION NUMBER: US/09/949,016

CURRENT FILING DATE: 2000-04-14

PRIOR APPLICATION NUMBER: 60/241,755

PRIOR FILING DATE: 2000-10-20

PRIOR APPLICATION NUMBER: 60/237,768

PRIOR FILING DATE: 2000-10-03

PRIOR APPLICATION NUMBER: 60/231,498

PRIOR FILING DATE: 2000-09-08

NUMBER OF SEQ ID NOS: 207012

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 41290

LENGTH: 601

TYPE: DNA

ORGANISM: Human

US-09-949-016-41290

Qy 1 AAAAAAAAAATCACGGCCAGGCATGG 25
|||
298 AAAAAAAAAAATCGGCCAGGCATGG 274
pb

RESULT 25
US-09-949-016-48718
; Sequence 48718, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 48718
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-48718

Qy

1 AAAAAAAAAATCAGGCCAGGCATGG 25
||||| |
pB

307 AAAAAAAAAAAAAAGGCCAGGCATGG 331
||||| |

RESULT 26
US-09-949-016-54603/c
; Sequence 54603, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASES, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 54603
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-54603

Qy 1 AAAAAAAAAATCACGGCCAGGCATGG 25
||| ||| | ||| ||| |||
Pb 360 AAAAAAAAAAAAAAAGGCCAGGCATGG 336

```

RESULT 27
US-09-949-016-54604/c
; Sequence 54604, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 54604
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-54604

```

Qy 1 AAAAAAAAAATCAGGCCAGGCATGG 25
|||
Db 137 AAAAAAAAAAAAAAGGCCAGGCATGG 113

; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-56302

Query Match 80.8%; Score 20.2; DB 3; Length 601;
Best Local Similarity 88.0%; Pred. No. 17;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAAAAAATCAGCGCCAGGCATGG 25
|||||
Db 594 AAAAAAAAAATCAGCGCCAGGCATGG 570

RESULT 37

US-09-949-016-59272/c
; Sequence 59272, Application US/09949016
; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; FILE REFERENCE: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 59272

; LENGTH: 601

; TYPE: DNA

; ORGANISM: Human

US-09-949-016-59272

Query Match 80.8%; Score 20.2; DB 3; Length 601;
Best Local Similarity 88.0%; Pred. No. 17;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAAAAAATCAGCGCCAGGCATGG 25
|||||
Db 586 AAAAAAAAAATCAGCGCCAGGCATGG 562

RESULT 38

US-09-949-016-63344/c
; Sequence 63344, Application US/09949016
; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; FILE REFERENCE: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 63344

; LENGTH: 601

; TYPE: DNA

; ORGANISM: Human

US-09-949-016-63344

Query Match 80.8%; Score 20.2; DB 3; Length 601;
Best Local Similarity 88.0%; Pred. No. 17;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAAAAAATCAGCGCCAGGCATGG 25
|||||
Db 415 AAAAAAAAAATCAGCGCCAGGCATGG 391

RESULT 39

US-09-949-016-119032/c
; Sequence 119032, Application US/09949016
; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; FILE REFERENCE: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 119032

; LENGTH: 601

; TYPE: DNA

; ORGANISM: Human

US-09-949-016-119032

Query Match 80.8%; Score 20.2; DB 3; Length 601;
Best Local Similarity 88.0%; Pred. No. 17;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAAAAAATCAGCGCCAGGCATGG 25
|||||
Db 360 AAAAAAAAAATCAGCGCCAGGCATGG 336

RESULT 40

US-09-949-016-119033/c
; Sequence 119033, Application US/09949016
; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; FILE REFERENCE: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 119033

; LENGTH: 601

; TYPE: DNA

; ORGANISM: Human

US-09-949-016-119033

Query Match 80.8%; Score 20.2; DB 3; Length 601;
Best Local Similarity 88.0%; Pred. No. 17;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAAAAAATCAGCGCCAGGCATGG 25

```
Db      137 AAAAAAAAAAAGGCCAGGCATGG 113
||||||| | |||||||||
RESULT 41
US-09-949-016-119034/c
; Sequence 119034, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 119034
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-119034
Query Match      80.8%; Score 20.2; DB 3; Length 601;
Best Local Similarity 88.0%; Pred. No. 17;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy      1 AAAAAAAAAATCAGGCCAGGCATGG 25
||||||| | |||||||||
Db      68 AAAAAAAAAAAGGCCAGGCATGG 44
||||||| | |||||||||

RESULT 42
US-09-949-016-119068/c
; Sequence 119068, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 119068
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-119068
Query Match      80.8%; Score 20.2; DB 3; Length 601;
Best Local Similarity 88.0%; Pred. No. 17;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy      1 AAAAAAAAAATCAGGCCAGGCATGG 25
||||||| | |||||||||
Db      360 AAAAAAAAAAAGGCCAGGCATGG 336
||||||| | |||||||||

RESULT 43
US-09-949-016-119069/c
; Sequence 119069, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 119069
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-119069
Query Match      80.8%; Score 20.2; DB 3; Length 601;
Best Local Similarity 88.0%; Pred. No. 17;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy      1 AAAAAAAAAATCAGGCCAGGCATGG 25
||||||| | |||||||||
Db      68 AAAAAAAAAAAGGCCAGGCATGG 44
||||||| | |||||||||

RESULT 45
US-09-949-016-119104/c
; Sequence 119104, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
```

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 119104
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-119104

Query Match 80.8%; Score 20.2; DB 3; Length 601;
Best Local Similarity 88.0%; Pred. No. 17;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAAAAAATCAGCGCCAGGCATGG 25
||||||| | |||||
Db 360 AAAAAAAAAAAGCCAGGCATGG 336

RESULT 46
US-09-949-016-119105/c
; Sequence 119105, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 119105
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-119105

Query Match 80.8%; Score 20.2; DB 3; Length 601;
Best Local Similarity 88.0%; Pred. No. 17;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAAAAAATCAGCGCCAGGCATGG 25
||||||| | |||||
Db 137 AAAAAAAAAAAGCCAGGCATGG 113

RESULT 47
US-09-949-016-119106/c
; Sequence 119106, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 119106
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-119106

Query Match 80.8%; Score 20.2; DB 3; Length 601;
Best Local Similarity 88.0%; Pred. No. 17;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAAAAAATCAGCGCCAGGCATGG 25
||||||| | |||||
Db 68 AAAAAAAAAAAGCCAGGCATGG 44

RESULT 48
US-09-949-016-119140/c
; Sequence 119140, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 119140
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-119140

Query Match 80.8%; Score 20.2; DB 3; Length 601;
Best Local Similarity 88.0%; Pred. No. 17;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAAAAAATCAGCGCCAGGCATGG 25
||||||| | |||||
Db 360 AAAAAAAAAAAGCCAGGCATGG 336

RESULT 49
US-09-949-016-119141/c
; Sequence 119141, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498

```

; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 119141
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-119141

Query Match          80.8%; Score 20.2; DB 3; Length 601;
Best Local Similarity 88.0%; Pred. No. 17;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AAAAAAAAAATCACGCCCAGGCATGG 25
    |||||  |||||  |||||  |||||
Db 137 AAAAAAAAAAGGCCAGGCATGG 113

RESULT 50
US-09-949-016-119142/c
; Sequence 119142, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 119142
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-119142

Query Match          80.8%; Score 20.2; DB 3; Length 601;
Best Local Similarity 88.0%; Pred. No. 17;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AAAAAAAAAATCACGCCCAGGCATGG 25
    |||||  |||||  |||||  |||||
Db 68 AAAAAAAAAAGGCCAGGCATGG 44

Search completed: February 3, 2006, 16:31:48
Job time : 86.6667 secs

```

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 3, 2006, 14:37:28 ; Search time 365.556 Seconds
(without alignments)
565.535 Million cell updates/sec

Title: US-10-719-900-1

Perfect score: 25

Sequence: 1 aaaaaaaaaacagccagcgatg 25

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database : Published Applications NA_Main:

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	25	100.0	25	8	US-10-719-900-1
2	23.4	93.6	25	8	US-10-719-900-2
3	23.4	93.6	41309	7	US-10-741-601-5771
4	21.8	87.2	504	4	US-09-925-065A-287766
5	21.8	87.2	510	3	US-09-796-692-4519
6	21.8	87.2	510	3	US-09-796-692-4519
7	21.8	87.2	510	3	US-10-040-862-4519
8	21.8	87.2	510	5	US-10-040-862-4519
9	21.8	87.2	510	5	US-10-057-475B-4519
10	21.8	87.2	510	6	US-10-057-475B-4519
11	21.8	87.2	510	6	US-10-154-884B-4519
12	21.8	87.2	510	6	US-10-154-884B-4519
13	21.8	87.2	510	8	US-10-764-324-4519
14	21.8	87.2	510	8	US-10-764-324-4519
15	21.8	87.2	516	3	US-09-796-692-4518
16	21.8	87.2	516	5	US-10-040-862-4518
17	21.8	87.2	516	6	US-10-057-475B-4518
18	21.8	87.2	516	6	US-10-154-884B-4518
19	21.8	87.2	516	8	US-10-764-324-4518
20	21.8	87.2	590	4	US-09-925-065A-46634
21	21.8	87.2	591	4	US-09-925-065A-677298
22	21.8	87.2	591	5	US-10-027-632-269686
23	21.8	87.2	591	6	US-10-027-632-269686

24	21.8	87.2	715	5	US-10-027-632-20441	Sequence 20441, A	
25	21.8	87.2	1870	4	US-10-027-632-20441	Sequence 20441, A	
c	26	21.8	715	6	US-09-925-065A-37155	Sequence 37155, A	
27	21.8	87.2	2969	4	US-09-925-065A-28421	Sequence 28421, A	
28	21.8	87.2	2969	4	US-09-925-065A-60862	Sequence 60862, A	
c	29	21.8	3067	6	US-10-057-475B-10476	Sequence 10476, A	
c	30	21.8	3067	6	US-10-154-884B-10476	Sequence 10476, A	
c	31	21.8	3105	6	US-10-057-475B-10477	Sequence 10477, A	
c	32	21.8	3105	6	US-10-154-884B-10477	Sequence 10477, A	
c	33	21.8	3288	6	US-10-144-194A-37	Sequence 37, Appl	
c	34	21.8	3288	8	US-10-491-566-37	Sequence 37, Appl	
c	35	21.8	51657	6	US-10-057-475B-10475	Sequence 10475, A	
c	36	21.8	51657	6	US-10-154-884B-10475	Sequence 10475, A	
c	37	21.8	98642	5	US-10-087-192-1105	Sequence 1105, Ap	
c	38	21.8	98642	5	US-10-087-192-1864	Sequence 1864, Ap	
c	39	20.8	1315	4	US-09-925-065A-39390	Sequence 39390, A	
c	40	20.8	25574	3	US-09-764-891-8307	Sequence 8307, Ap	
c	41	20.8	25576	3	US-09-764-891-8306	Sequence 8306, Ap	
c	42	20.8	27893	6	US-10-017-161-757	Sequence 757, App	
c	43	20.8	27893	5	US-10-087-192-1630	Sequence 1630, Ap	
c	44	20.8	599001	9	US-10-317-869A-4	Sequence 4, Appli	
c	45	20.4	201	7	US-10-741-601-24054	Sequence 24054, A	
c	46	20.4	201	8	US-10-741-600-64920	Sequence 64920, A	
c	47	20.4	528	4	US-09-925-065A-163238	Sequence 163238, A	
c	48	20.4	528	5	US-10-027-632-201503	Sequence 201503, A	
c	49	20.4	546	6	US-10-027-632-201503	Sequence 201503, A	
c	50	20.4	573	4	US-09-925-065A-51348	Sequence 51348, A	
c	51	20.4	590	4	US-09-925-065A-323966	Sequence 323966, A	
c	52	20.4	1511	4	US-09-925-065A-681763	Sequence 681763, A	
c	53	20.4	15804	7	US-10-741-601-56633	Sequence 56633, Ap	
c	54	20.4	15804	8	US-10-741-600-17690	Sequence 17690, Ap	
c	55	20.4	24446	7	US-10-741-601-5740	Sequence 5740, Ap	
c	56	20.4	24446	8	US-10-741-600-17889	Sequence 17889, A	
c	57	20.4	41907	3	US-09-967-013-5	Sequence 5, Appli	
c	58	20.4	210204	8	US-10-723-860-1746	Sequence 1746, Ap	
c	59	20.4	210204	9	US-10-756-149-1685	Sequence 1685, Ap	
c	60	20.2	80.8	201	8	US-10-741-600-34716	Sequence 34716, A
c	61	20.2	80.8	344	3	US-09-803-719-718	Sequence 718, App
c	62	20.2	80.8	344	9	US-10-779-543-13988	Sequence 13988, A
c	63	20.2	80.8	384	3	US-09-995-793A-14	Sequence 14, Appl
c	64	20.2	80.8	440	4	US-09-925-065A-917840	Sequence 917840, A
c	65	20.2	80.8	440	4	US-09-925-065A-944841	Sequence 944841, A
c	66	20.2	80.8	463	8	US-10-674-124A-18846	Sequence 18846, A
c	67	20.2	80.8	465	3	US-09-867-701-4978	Sequence 4978, Ap
c	68	20.2	80.8	471	4	US-09-925-065A-232436	Sequence 232436, A
c	69	20.2	80.8	471	4	US-09-925-065A-232437	Sequence 232437, A
c	70	20.2	80.8	471	4	US-09-925-065A-232438	Sequence 232438, A
c	71	20.2	80.8	480	9	US-10-779-543-17707	Sequence 17707, A
c	72	20.2	80.8	484	3	US-09-864-761-11717	Sequence 11717, A
c	73	20.2	80.8	514	4	US-09-925-065A-599935	Sequence 599935, A
c	74	20.2	80.8	514	4	US-09-925-065A-599936	Sequence 599936, A
c	75	20.2	80.8	516	5	US-10-027-632-289481	Sequence 289481, A
c	76	20.2	80.8	516	6	US-10-027-632-289481	Sequence 289481, A
c	77	20.2	80.8	526	9	US-10-779-543-20631	Sequence 20631, A
c	78	20.2	80.8	530	4	US-09-925-065A-939463	Sequence 939463, A
c	79	20.2	80.8	530	4	US-09-925-065A-939464	Sequence 939464, A
c	80	20.2	80.8	537	5	US-10-027-632-78934	Sequence 78934, A
c	81	20.2	80.8	537	5	US-10-027-632-78935	Sequence 78935, A
c	82	20.2	80.8	537	5	US-10-027-632-306937	Sequence 306937, A
c	83	20.2	80.8	537	5	US-10-027-632-306938	Sequence 306938, A
c	84	20.2	80.8	537	6	US-10-027-632-78934	Sequence 78934, A
c	85	20.2	80.8	537	6	US-10-027-632-78935	Sequence 78935, A
c	86	20.2	80.8	537	6	US-10-027-632-306937	Sequence 306937, A
c	87	20.2	80.8	537	6	US-10-027-632-306938	Sequence 306938, A
c	88	20.2	80.8	543	4	US-09-925-065A-17381	Sequence 17381, A
c	89	20.2	80.8	556	4	US-09-925-065A-906458	Sequence 906458, A
c	90	20.2	80.8	564	4	US-09-925-065A-580000	Sequence 580000, A
c	91	20.2	80.8	570	5	US-10-027-632-24416	Sequence 24416, A
c	92	20.2	80.8	570	5	US-10-027-632-24416	Sequence 24416, A
c	93	20.2	80.8	573	5	US-10-027-632-277986	Sequence 277986, A
c	94	20.2	80.8	573	6	US-10-027-632-277986	Sequence 277986, A
c	95	20.2	80.8	599	4	US-09-925-065A-662607	Sequence 662607, A
c	96	20.2	80.8	606	4	US-09-925-065A-770354	Sequence 770354, A

```

97 20.2 80.8 609 4 US-09-925-065A-105459 Sequence 105459,
98 20.2 80.8 609 4 US-09-925-065A-105460 Sequence 105460,
99 20.2 80.8 630 4 US-09-925-065A-937990 Sequence 937990,
100 20.2 80.8 630 5 US-10-027-632-129451 Sequence 129451,
101 20.2 80.8 631 6 US-10-027-632-129451 Sequence 129451,
102 20.2 80.8 631 4 US-09-925-065A-940220 Sequence 940220,
103 20.2 80.8 631 4 US-09-925-065A-923002 Sequence 923002,
104 20.2 80.8 640 4 US-09-925-065A-53378 Sequence 53378, A
105 20.2 80.8 645 4 US-09-925-065A-732668 Sequence 732668,
106 20.2 80.8 645 4 US-09-925-065A-732669 Sequence 732669,
107 20.2 80.8 703 4 US-09-925-065A-940164 Sequence 940164,
108 20.2 80.8 703 4 US-09-925-065A-940165 Sequence 940165,
109 20.2 80.8 703 4 US-09-925-065A-955134 Sequence 955134,
110 20.2 80.8 728 4 US-09-925-065A-21545 Sequence 21545, A
111 20.2 80.8 728 4 US-09-925-065A-21546 Sequence 21546, A
112 20.2 80.8 771 5 US-10-027-632-150013 Sequence 150013,
113 20.2 80.8 771 6 US-10-027-632-150013 Sequence 150013,
114 20.2 80.8 789 4 US-09-925-065A-953148 Sequence 953148,
115 20.2 80.8 795 5 US-10-027-632-164373 Sequence 164373,
116 20.2 80.8 795 5 US-10-027-632-164373 Sequence 164373,
117 20.2 80.8 841 5 US-10-027-632-161160 Sequence 161160,
118 20.2 80.8 841 6 US-10-027-632-161160 Sequence 161160,
119 20.2 80.8 855 5 US-10-027-632-29404 Sequence 29404, A
120 20.2 80.8 856 6 US-10-027-632-162444 Sequence 162444,
121 20.2 80.8 856 5 US-10-027-632-162445 Sequence 162445,
122 20.2 80.8 856 5 US-10-027-632-162445 Sequence 162445,
123 20.2 80.8 856 6 US-10-027-632-162444 Sequence 162444,
124 20.2 80.8 856 6 US-10-027-632-162445 Sequence 162445,
125 20.2 80.8 957 5 US-10-027-632-258749 Sequence 258749,
126 20.2 80.8 957 5 US-10-027-632-258750 Sequence 258750,
127 20.2 80.8 957 6 US-10-027-632-258749 Sequence 258749,
128 20.2 80.8 957 6 US-10-027-632-258750 Sequence 258750,
129 20.2 80.8 974 4 US-09-925-065A-54818 Sequence 54818, A
130 20.2 80.8 1012 4 US-09-925-065A-17445 Sequence 17445, A
131 20.2 80.8 1054 4 US-09-925-065A-237766 Sequence 237766,
132 20.2 80.8 1103 5 US-10-027-632-30415 Sequence 30415, A
133 20.2 80.8 1103 6 US-10-027-632-30415 Sequence 30415, A
134 20.2 80.8 1131 3 US-09-764-878-260 Sequence 260, App
135 20.2 80.8 1131 3 US-09-764-860-939 Sequence 939, App
136 20.2 80.8 1131 5 US-10-079-854-260 Sequence 260, App
137 20.2 80.8 1131 5 US-10-074-095-939 Sequence 939, App
138 20.2 80.8 1131 6 US-10-212-872-939 Sequence 939, App
139 20.2 80.8 1140 5 US-10-027-632-254873 Sequence 254873,
140 20.2 80.8 1140 6 US-10-027-632-254873 Sequence 254873,
141 20.2 80.8 1264 5 US-10-001-857-59 Sequence 59, Appl
142 20.2 80.8 1507 4 US-09-925-065A-700674 Sequence 700674,
143 20.2 80.8 1512 4 US-09-925-065A-680429 Sequence 680429,
144 20.2 80.8 1512 4 US-09-925-065A-680430 Sequence 680430,
145 20.2 80.8 1512 4 US-09-925-065A-680431 Sequence 680431,
146 20.2 80.8 1981 4 US-09-925-065A-717860 Sequence 717860,
147 20.2 80.8 1981 4 US-09-925-065A-717861 Sequence 717861,
148 20.2 80.8 2021 4 US-09-925-065A-27392 Sequence 27392, A
149 20.2 80.8 6186 3 US-09-764-860-1075 Sequence 1075, Ap
150 20.2 80.8 6186 5 US-10-074-095-1075 Sequence 1075, Ap

ALIGNMENTS

RESULT 1
US-10-719-900-1
; Sequence 1, Application US/10719900
; Publication No. US20050026164A1
; GENERAL INFORMATION:
; APPLICANT: Xue Mei Zhou
; TITLE OF INVENTION: Methods of Genetic Analysis of Mouse
; FILE REFERENCE: 3528.1
; CURRENT FILING DATE: 2003-11-20
; PRIOR APPLICATION NUMBER: 60/427,808
; PRIOR FILING DATE: 2002 11 20
; NUMBER OF SEQ ID NOS: 982914
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1

Query Match 100.0%; Score 25; DB 8; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.14;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAAAAAAATCAGCGCCAGGCATGG 25
Db 1 AAAAAAAAAATCAGCGCCAGGCATGG 25

RESULT 2
US-10-719-900-2
; Sequence 2, Application US/10719900
; Publication No. US20050026164A1
; GENERAL INFORMATION:
; APPLICANT: Xue Mei Zhou
; TITLE OF INVENTION: Methods of Genetic Analysis of Mouse
; FILE REFERENCE: 3528.1
; CURRENT FILING DATE: 2003-11-20
; PRIOR APPLICATION NUMBER: 60/427,808
; PRIOR FILING DATE: 2002 11 20
; NUMBER OF SEQ ID NOS: 982914
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 2
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-719-900-2
Query Match 93.6%; Score 23.4; DB 8; Length 25;
Best Local Similarity 96.0%; Pred. No. 0.74;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAAAAAAAATCAGCGCCAGGCATGG 25
Db 1 AAAAAAAAAATCAGCGCCAGGCATGG 25

RESULT 3
US-10-741-601-5771/c
; Sequence 5771, Application US/10741601
; Publication No. US20040166519A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001500
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 26415
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5771
; LENGTH: 41309
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-741-601-5771
Query Match 93.6%; Score 23.4; DB 7; Length 41309;
Best Local Similarity 96.0%; Pred. No. 2.3;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAAAAAAAATCAGCGCCAGGCATGG 25
Db 19143 AAAAAAAAAATCAGCGCCAGGCATGG 19119

RESULT 4
```


US-09-925-065A-287766
; Sequence 287766, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 287766
; LENGTH: 504
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-287766

Query Match 87.2%; Score 21.8; DB 4; Length 504;
Best Local Similarity 92.0%; Pred. No. 6.5;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAAAAAACAGCGCCAGGCATGG 25
||||| ||||| ||||| ||||| |||||
DB 329 AAAAAAAAAATTAAGCGCAGGCATGG 353

RESULT 5
US-09-796-692-4519
; Sequence 4519, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Mannion, Paul A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 9149
; LENGTH: 510
; TYPE: DNA

; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4519
; LENGTH: 510
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (28)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (242)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (344)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (472)
; OTHER INFORMATION: n=A,T,C or G
US-09-796-692-4519

Query Match 87.2%; Score 21.8; DB 3; Length 510;
Best Local Similarity 92.0%; Pred. No. 6.5;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAAAAAACAGCGCCAGGCATGG 25
||||| ||||| ||||| ||||| |||||
DB 405 AAAAAAAAAATTAAGCGCAGGCATGG 429

RESULT 6
US-09-796-692-9149
; Sequence 9149, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Mannion, Paul A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 9149
; LENGTH: 510
; TYPE: DNA

OTHER INFORMATION: n=A,T,C or G
FEATURE: unsure
NAME/KEY: (481)
LOCATION: (481)
OTHER INFORMATION: n=A,T,C or G
FEATURE: unsure
NAME/KEY: (504)
LOCATION: (504)
OTHER INFORMATION: n=A,T,C or G
US-10-040-862-9149

Query Match 87.2%; Score 21.8; DB 5; Length 510;
Best Local Similarity 92.0%; Pred. No. 6.5;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AAAAAAAAAATCAGCGCCAGGCATGG 25
||||| ||||||| ||||||| ||||||| ||
Db 405 AAAAAATAATCAGCGCCAGGCACGG 429

RESULT 9
US-10-057-475B-4519
; Sequence 4519, Application US/10057475B
; Publication No. US2004002068A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Mannion, Jane
; APPLICANT: Clapper, Jonathan David
; APPLICANT: Wang, Aijun
; APPLICANT: Ordonez, Nadia
; APPLICANT: Carter, Lauren
; APPLICANT: McNeill, Patricia Dianne
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; TITLE OF INVENTION: Hematological Malignancies
; FILE REFERENCE: 014058-014402US
; CURRENT APPLICATION NUMBER: US/10/057,475B
; CURRENT FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 10979
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 4519
; LENGTH: 510
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(510)
; OTHER INFORMATION: n = G, A, C or T
US-10-057-475B-4519

Query Match 87.2%; Score 21.8; DB 6; Length 510;

Best Local Similarity 92.0%; Pred. No. 6.5;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 AAAAAAAAAATCAGCGCCAGGCATGG 25
||||| ||||||| ||||||| ||||||| ||
Db 405 AAAAAATAATCAGCGCCAGGCACGG 429

RESULT 10
US-10-057-475B-9149
; Sequence 9149, Application US/10057475B
; Publication No. US2004002068A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Clapper, Jonathan David
; APPLICANT: Wang, Aijun
; APPLICANT: Ordonez, Nadia
; APPLICANT: Carter, Lauren
; APPLICANT: McNeill, Patricia Dianne
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; TITLE OF INVENTION: Hematological Malignancies
; FILE REFERENCE: 014058-014402US
; CURRENT APPLICATION NUMBER: US/10/057,475B
; CURRENT FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 10979
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 9149
; LENGTH: 510
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(510)
; OTHER INFORMATION: n = G, A, C or T
US-10-057-475B-9149

Query Match 87.2%; Score 21.8; DB 6; Length 510;
Best Local Similarity 92.0%; Pred. No. 6.5;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AAAAAAAAAATCAGCGCCAGGCATGG 25
||||| ||||||| ||||||| ||||||| ||
Db 405 AAAAAATAATCAGCGCCAGGCACGG 429

RESULT 11
US-10-154-884B-4519
; Sequence 4519, Application US/10154884B
; Publication No. US20040005561A1

```

; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 11290
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 9149
; LENGTH: 510
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(510)
; OTHER INFORMATION: n = g, a, c or t
US-10-154-884B-9149

Query Match      87.3%   Score 21.8; DB 6; Length 510;
Best Local Similarity 92.08; Pred. No. 6.5;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0

QY 1 AAAAAAAAAATCAGGCCAGGCATGG 25
   ||||| ||||| ||||| ||||| |||||
DB 405 AAAAAATAAATCAGGCCAGGCACGG 429

RESULT 13
US-10-764-324-4519
; Sequence 4519, Application US/10764324
; Publication No. US2004017539A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and
; FILE REFERENCE: 014058-013520US
; CURRENT APPLICATION NUMBER: US/10764,324
; CURRENT FILING DATE: 2004-01-23
; PRIOR APPLICATION NUMBER: US/10/040,862
; PRIOR FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; Remaining Prior Application data removed - See File Wrapper or PALM.

```

```
;
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4519
; LENGTH: 510
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (28)
; OTHER INFORMATION: n=A,T,C or G
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (242)
; OTHER INFORMATION: n=A,T,C or G
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (344)
; OTHER INFORMATION: n=A,T,C or G
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (472)
; OTHER INFORMATION: n=A,T,C or G
; US-10-764-324-4519
```

```
Query Match 87.2%; Score 21.8; DB 8; Length 510;
Best Local Similarity 92.0%; Pred. No. 6.5;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy 1 AAAAAAAAAATCAGCGCCAGGCATGG 25
    ||||| ||||| ||||| ||||| |||||
Db 405 AAAAAATAATCAGCGCCAGGCACGG 429
```

```
RESULT 14
US-10-764-324-9149
; Sequence 9149, Application US/10764324
; Publication No. US2004017539A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Mannion, Jane
; APPLICANT: Algate, Paul A.
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; TITLE OF INVENTION: Hematological Malignancies
; FILE REFERENCE: 014058-013520US
; CURRENT APPLICATION NUMBER: US/10/764,324
; CURRENT FILING DATE: 2004-01-23
; PRIOR APPLICATION NUMBER: US/10/040,862
; PRIOR FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 9149
; LENGTH: 510
```

```
;
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (472)
; OTHER INFORMATION: n=A,T,C or G
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (481)
; OTHER INFORMATION: n=A,T,C or G
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (504)
; OTHER INFORMATION: n=A,T,C or G
; US-10-764-324-9149
```

```
Query Match 87.2%; Score 21.8; DB 8; Length 510;
Best Local Similarity 92.0%; Pred. No. 6.5;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy 1 AAAAAAAAAATCAGCGCCAGGCATGG 25
    ||||| ||||| ||||| ||||| |||||
Db 405 AAAAAATAATCAGCGCCAGGCACGG 429
```

```
RESULT 15
US-09-796-692-4518
; Sequence 4518, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Mannion, Jane
; APPLICANT: Algate, Paul A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAP
; TITLE OF INVENTION: HEMATOLOGICAL MALIGNANCIES
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4518
; LENGTH: 516
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-796-692-4518
```

```
Query Match 87.2%; Score 21.8; DB 3; Length 516;
Best Local Similarity 92.0%; Pred. No. 6.5;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

QY 1 AAAAAAAAAATCAGCGCCAGGCATGG 25
||||| ||||||| ||||| |||
Db 405 AAAAAATAATCAGCGCCAGGCACGG 429

RESULT 16
US-10-040-862-4518
; Sequence 4518, Application US/10040862
; Publication No. US20030078396A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; TITLE OF INVENTION: Hematological Malignancies
; FILE REFERENCE: 014058-013520US
; CURRENT APPLICATION NUMBER: US/10/040,862
; CURRENT FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 60/223,378
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: US 09/796,692
; PRIOR FILING DATE: 2001-03-01
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4518
; LENGTH: 516
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-040-862-4518

Query Match 87.2%; Score 21.8; DB 5; Length 516;
Best Local Similarity 92.0%; Pred. No. 6.5;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAAAAAATCAGCGCCAGGCATGG 25
||||| ||||||| ||||| |||
Db 405 AAAAAATAATCAGCGCCAGGCACGG 429

RESULT 17
US-10-057-475B-4518
; Sequence 4518, Application US/10057475B
; Publication No. US2004002068A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Clapper, Jonathan David

; APPLICANT: Wang, Aijun
; APPLICANT: Ordenez, Nadia
; APPLICANT: Carter, Lauren
; APPLICANT: McNeill, Patricia Dianne
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; TITLE OF INVENTION: Hematological Malignancies
; FILE REFERENCE: 014058-014402US
; CURRENT APPLICATION NUMBER: US/10/057,475B
; CURRENT FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 10979
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4518
; LENGTH: 516
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-057-475B-4518

Query Match 87.2%; Score 21.8; DB 6; Length 516;
Best Local Similarity 92.0%; Pred. No. 6.5;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAAAAAATCAGCGCCAGGCATGG 25
||||| ||||||| ||||| |||
Db 405 AAAAAATAATCAGCGCCAGGCACGG 429

RESULT 18
US-10-154-884B-4518
; Sequence 4518, Application US/10154884B
; Publication No. US20040005561A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc W.
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; TITLE OF INVENTION: Hematological Malignancies
; FILE REFERENCE: 014058-013521US
; CURRENT APPLICATION NUMBER: US/10/154,884B
; CURRENT FILING DATE: 2002-05-23
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28

```
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 11290
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4518
; LENGTH: 516
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-154-884B-4518

Query Match      87.2%; Score 21.8; DB 6; Length 516;
Best Local Similarity 92.0%; Pred. No. 6.5;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 AAAAAAAAAATCACGCCGAGCATGG 25
      ||||| ||||| ||||| ||||| |||||
Db      405 AAAAAATTAATCACGCCGAGCATGG 429

RESULT 19
US-10-764-324-4518
; Sequence 4518, Application US/10764324
; Publication No. US2004017539A1
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; TITLE OF INVENTION: Hematological Malignancies
; FILE REFERENCE: 014058-013520US
; CURRENT APPLICATION NUMBER: US/10/764,324
; CURRENT FILING DATE: 2004-01-23
; PRIOR APPLICATION NUMBER: US/10/040,862
; PRIOR FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4518
; LENGTH: 516
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-764-324-4518

Query Match      87.2%; Score 21.8; DB 8; Length 516;
```

```
Best Local Similarity 92.0%; Pred. No. 6.5;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 AAAAAAAAAATCACGCCGAGCATGG 25
      ||||| ||||| ||||| ||||| |||||
Db      405 AAAAAATTAATCACGCCGAGCATGG 429

RESULT 20
US-09-925-065A-46634/c
; Sequence 46634, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 46634
; LENGTH: 590
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-46634

Query Match      87.2%; Score 21.8; DB 4; Length 590;
Best Local Similarity 92.0%; Pred. No. 6.6;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 AAAAAAAAAATCACGCCGAGCATGG 25
      ||||| ||||| ||||| ||||| |||||
Db      129 AAAAAAAATCTTGGCCAGCATGG 105

RESULT 21
US-09-925-065A-677298
; Sequence 677298, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 677298
; LENGTH: 591
; TYPE: DNA
; ORGANISM: Homo sapiens
```

US-09-925-065A-677298

Query Match 87.2%; Score 21.8; DB 4; Length 591;
Best Local Similarity 92.0%; Pred. No. 6.6;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAAAAAATCAGGCCAGGCATGG 25
||||| ||| ||||| ||||| |||||
Db 456 AAAAAAAAAATCAGGCCAGGCATGG 480

RESULT 22

US-10-027-632-269686
; Sequence 269686, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129

; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 269686
; LENGTH: 591
; TYPE: DNA
; ORGANISM: Human

US-10-027-632-269686

Query Match 87.2%; Score 21.8; DB 5; Length 591;
Best Local Similarity 92.0%; Pred. No. 6.6;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAAAAAATCAGGCCAGGCATGG 25
||||| ||| ||||| ||||| |||||
Db 64 AAAAAAAAAATCAGGCCAGGCACGG 88

RESULT 23

US-10-027-632-269686
; Sequence 269686, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129

; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363

; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 269686
; LENGTH: 591
; TYPE: DNA
; ORGANISM: Human

US-10-027-632-269686

Query Match 87.2%; Score 21.8; DB 6; Length 591;
Best Local Similarity 92.0%; Pred. No. 6.6;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAAAAAATCAGGCCAGGCATGG 25
||||| ||| ||||| ||||| |||||
Db 64 AAAAAAAAAATCAGGCCAGGCACGG 88

RESULT 24

US-10-027-632-20441
; Sequence 20441, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129

; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20441
; LENGTH: 715
; TYPE: DNA
; ORGANISM: Human

US-10-027-632-20441

Query Match 87.2%; Score 21.8; DB 5; Length 715;
Best Local Similarity 92.0%; Pred. No. 6.8;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAAAAAATCAGGCCAGGCATGG 25
||||| ||| ||||| ||||| |||||
Db 402 AAAAAAAAAATCAGGCCAGGCACGG 426

RESULT 25

US-10-027-632-20441
; Sequence 20441, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129

; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 20441
; LENGTH: 715
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-20441

Query Match 87.2%; Score 21.8; DB 6; Length 715;
Best Local Similarity 92.0%; Pred. No. 6.8;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AAAAAAAAAATCAGCGCCAGCGCATGG 25
||||| ||||||| ||||||| ||||||| ||
Db 402 AAAAAATAAATCAGCGCCAGCGCATGG 426

RESULT 26

US-09-925-065A-37155/c
; Sequence 37155, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 37155
; LENGTH: 1870
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-37155

Query Match 87.2%; Score 21.8; DB 4; Length 1870;
Best Local Similarity 92.0%; Pred. No. 7.9;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AAAAAAAAAATCAGCGCCAGCGCATGG 25
||||| ||||||| ||||||| ||||||| ||
Db 710 AAAAAATAAATCAGCGCCAGCGCATGG 686

RESULT 27

US-09-925-065A-28421

; Sequence 28421, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 28421
; LENGTH: 2969
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-28421

Query Match 87.2%; Score 21.8; DB 4; Length 2969;
Best Local Similarity 92.0%; Pred. No. 8.5;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AAAAAAAAAATCAGCGCCAGCGCATGG 25
||||| ||||||| ||||||| ||||||| ||
Db 1239 AAAAAATAATCATGGCCAGCGGTGG 1263

RESULT 28

US-09-925-065A-60862
; Sequence 60862, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 60862
; LENGTH: 2969
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-60862

Query Match 87.2%; Score 21.8; DB 4; Length 2969;
Best Local Similarity 92.0%; Pred. No. 8.5;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AAAAAAAAAATCAGCGCCAGCGCATGG 25
||||| ||||||| ||||||| ||||||| ||
Db 1239 AAAAAATAATCATGGCCAGCGGTGG 1263

```
RESULT 29
US-10-057-475B-10476/c
; Sequence 10476, Application US/10057475B
; Publication No. US20040002068A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Clapper, Jonathan David
; APPLICANT: Wang, Aijun
; APPLICANT: Ordonez, Nadia
; APPLICANT: Carter, Lauren
; APPLICANT: McNeill, Patricia Dianne
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; TITLE OF INVENTION: Hematological Malignancies
; FILE REFERENCE: 014058-014402US
; CURRENT FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: US/10/057,475B
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/222,903
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 11290
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10476
; LENGTH: 3067
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-057-475B-10476

Query Match      87.2%; Score 21.8; DB 6; Length 3067;
Best Local Similarity 92.0%; Pred. No. 8.6;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAAAAAATCAGCGCCAGCGCATGG 25
DB 1184 AAAAAATAATCAGCGCCAGCGCATGG 1160

RESULT 30
US-10-057-475B-10476/c
; Sequence 10476, Application US/10154884B
; Publication No. US20040005561A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc W.
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; TITLE OF INVENTION: Hematological Malignancies
; FILE REFERENCE: 014058-013521US
; CURRENT FILING DATE: 2002-05-23
; CURRENT APPLICATION NUMBER: US/10/154,884B
```

```
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 11290
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10476
; LENGTH: 3067
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-154-884B-10476

Query Match      87.2%; Score 21.8; DB 6; Length 3067;
Best Local Similarity 92.0%; Pred. No. 8.6;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAAAAAATCAGCGCCAGCGCATGG 25
DB 1184 AAAAAATAATCAGCGCCAGCGCATGG 1160

RESULT 31
US-10-057-475B-10477/c
; Sequence 10477, Application US/10057475B
; Publication No. US20040002068A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Clapper, Jonathan David
; APPLICANT: Wang, Aijun
; APPLICANT: Ordonez, Nadia
; APPLICANT: Carter, Lauren
; APPLICANT: McNeill, Patricia Dianne
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; TITLE OF INVENTION: Hematological Malignancies
; FILE REFERENCE: 014058-014402US
; CURRENT APPLICATION NUMBER: US/10/057,475B
; CURRENT FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
```

; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 10979
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10477
; LENGTH: 3105
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-057-475B-10477

Query Match 87.2%; Score 21.8; DB 6; Length 3105;
Best Local Similarity 92.0%; Pred. No. 8.6;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AAAAAAAAAATCAGCCGAGGATGG 25
||||| ||||||| ||||||| |||||
Db 1177 AAAAAATAATCAGCCGAGGATGG 1153

RESULT 32
US-10-154-884B-10477/c
; Sequence 10477, Application US/10154884B
; Publication No. US20040005561A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc W.
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; TITLE OF INVENTION: Hematological Malignancies
; FILE REFERENCE: 014058-013521US
; CURRENT APPLICATION NUMBER: US/10/154,884B
; CURRENT FILING DATE: 2002-05-23
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 11290
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10477
; LENGTH: 3105
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-154-884B-10477

Query Match 87.2%; Score 21.8; DB 6; Length 3105;
Best Local Similarity 92.0%; Pred. No. 8.6;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AAAAAAAAAATCAGCCGAGGATGG 25
||||| ||||||| ||||||| |||||
Db 1177 AAAAAATAATCAGCCGAGGATGG 1153

RESULT 33
US-10-144-194A-37
; Sequence 37, Application US/10144194A
; Publication No. US20030215809A1
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies Inc
; TITLE OF INVENTION: Regulated Breast Cancer Genes
; FILE REFERENCE: 3U 103 R1
; CURRENT APPLICATION NUMBER: US/10/144,194A
; CURRENT FILING DATE: 2002-06-12
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 37
; LENGTH: 3288
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (267)..(653)
US-10-144-194A-37

Query Match 87.2%; Score 21.8; DB 6; Length 3288;
Best Local Similarity 92.0%; Pred. No. 8.7;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AAAAAAAAAATCAGCCGAGGATGG 25
||||| ||||||| ||||||| |||||
Db 2979 AAAAAATAATCTTGGCCAGGATGG 3003

RESULT 34
US-10-491-566-37
; Sequence 37, Application US/10491566
; Publication No. US20040249144A1
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies Inc
; TITLE OF INVENTION: Regulated Breast Cancer Genes
; FILE REFERENCE: 3U 103 R1
; CURRENT APPLICATION NUMBER: US/10/491,566
; CURRENT FILING DATE: 2004-04-05
; NUMBER OF SEQ ID NOS: 148
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 37
; LENGTH: 3288
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (267)..(653)
; OTHER INFORMATION:
US-10-491-566-37

Query Match 87.2%; Score 21.8; DB 8; Length 3288;
Best Local Similarity 92.0%; Pred. No. 8.7;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AAAAAAAAAATCAGCCGAGGATGG 25
||||| ||||||| ||||||| |||||
Db 2979 AAAAAATAATCTTGGCCAGGATGG 3003

RESULT 35
US-10-057-475B-10475/c
; Sequence 10475, Application US/10057475B
; Publication No. US20040002068A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Clapper, Jonathan David
; APPLICANT: Wang, Aijun

```

: APPLICANT: Ordonez, Nadia
: APPLICANT: Carter, Lauren
: APPLICANT: McNeill, Patricia Dianne
: APPLICANT: Corixa Corporation
: TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
: TITLE OF INVENTION: Hematological Malignancies
: FILE REFERENCE: 014058-014402US
: CURRENT APPLICATION NUMBER: US/10/057,475B
: CURRENT FILING DATE: 2002-01-22
: PRIOR APPLICATION NUMBER: US 60/186,126
: PRIOR FILING DATE: 2000-03-01
: PRIOR APPLICATION NUMBER: US 60/190,479
: PRIOR FILING DATE: 2000-03-17
: PRIOR APPLICATION NUMBER: US 60/200,545
: PRIOR FILING DATE: 2000-04-27
: PRIOR APPLICATION NUMBER: US 60/200,303
: PRIOR FILING DATE: 2000-04-28
: PRIOR APPLICATION NUMBER: US 60/200,779
: PRIOR FILING DATE: 2000-04-28
: PRIOR APPLICATION NUMBER: US 60/200,999
: PRIOR FILING DATE: 2000-05-01
: PRIOR APPLICATION NUMBER: US 60/202,084
: PRIOR FILING DATE: 2000-05-04
: PRIOR APPLICATION NUMBER: US 60/206,201
: PRIOR FILING DATE: 2000-05-22
: PRIOR APPLICATION NUMBER: US 60/218,950
: PRIOR FILING DATE: 2000-07-14
: PRIOR APPLICATION NUMBER: US 60/222,903
: PRIOR FILING DATE: 2000-08-03
: Remaining Prior Application data removed - See File Wrapper or PALM.
: NUMBER OF SEQ ID NOS: 10979
: SOFTWARE: FastSEQ for Windows Version 3.0
: SEQ ID NO 10475
: LENGTH: 51657
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc feature
: LOCATION: (1)..(51657)
: OTHER INFORMATION: n = g, a, c or t
:
US-10-057-475B-10475

Query Match      87.2%; Score 21.8; DB 6; Length 51657;
Best Local Similarity 92.0%; Pred. No. 13;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AAAAAAAAAACAGCGCCAGGCATGG 25
   ||||| ||||| ||||| ||||| |||||
Db 28204 AAAAAATAATCACGCGCCAGGCATGG 28180

RESULT 36
US-10-154-884B-10475/c
: Sequence 10475, Application US/10154884B
: Publication No. US20040005561A1
: GENERAL INFORMATION:
: APPLICANT: Gaiger, Alexander
: APPLICANT: Algate, Paul A.
: APPLICANT: Mannion, Jane
: APPLICANT: Retter, Marc W.
: APPLICANT: Corixa Corporation
: TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
: TITLE OF INVENTION: Hematological Malignancies
: FILE REFERENCE: 014058-013521US
: CURRENT APPLICATION NUMBER: US/10/154,884B
: CURRENT FILING DATE: 2002-05-23
: PRIOR APPLICATION NUMBER: US 60/186,126
: PRIOR FILING DATE: 2000-03-01
: PRIOR APPLICATION NUMBER: US 60/190,479
: PRIOR FILING DATE: 2000-03-17
: PRIOR APPLICATION NUMBER: US 60/200,545
: PRIOR FILING DATE: 2000-04-27
: PRIOR APPLICATION NUMBER: US 60/200,303

```

```

US-10-087-192-1864/c
; Publication No. US20030077808A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8307
; LENGTH: 25574
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-891-8307

Query Match      83.2%; Score 20.8; DB 3; Length 25574;
Best Local Similarity 91.7%; Pred. No. 34;
Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAAAAAATCAGGCCAGGCATG 24
   |||||
Db 12227 AAAAAAAAAATCAAGGCCAGGCATG 12250

RESULT 41
US-09-764-891-8306
; Sequence 8306, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8306
; LENGTH: 25576
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-891-8306

Query Match      83.2%; Score 20.8; DB 3; Length 25576;
Best Local Similarity 91.7%; Pred. No. 34;
Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAAAAAATCAGGCCAGGCATG 24
   |||||
Db 12225 AAAAAAAAAATCAAGGCCAGGCATG 12248

RESULT 42
US-10-017-161-757
; Sequence 757, Application US/10017161
; Publication No. US20030143668A1
; GENERAL INFORMATION:
; APPLICANT: SUWA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: ABURATANI, HIROYUKI
; TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 084335/0152
; CURRENT APPLICATION NUMBER: US/10/017,161
; CURRENT FILING DATE: 2002-12-18
; PRIOR FILING DATE: 2001-06-18
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 2430
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 757
; LENGTH: 27893
; TYPE: DNA

```



```
Query Match      83.2%; Score 20.8; DB 9; Length 599001;
Best Local Similarity 91.7%; Pred. No. 56;
Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AAAAAAAAAATCAGCGCCAGGATGG 25
   |||||
Db 330645 AAAAAAAAAATCAGCGCCAGGATGG 330622

RESULT 45
US-10-741-601-24054/c
; Sequence 24054, Application US/10741601
; Publication No. US20040166519A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: STENOSIS, METHODS OF DETECTION AND USES THEREOF
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 26415
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 24054
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-741-601-24054

Query Match      81.6%; Score 20.4; DB 7; Length 201;
Best Local Similarity 95.5%; Pred. No. 25;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAAAAAAAATCAGCGCCAGGCA 22
   |||||
Db 109 AAAAAAAAAATCAGCGCCAGGCA 88

RESULT 46
US-10-741-600-64920/c
; Sequence 64920, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 64920
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-741-600-64920

Query Match      81.6%; Score 20.4; DB 8; Length 201;
Best Local Similarity 95.5%; Pred. No. 25;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAAAAAAAATCAGCGCCAGGCA 22
   |||||
Db 109 AAAAAAAAAATCAGCGCCAGGCA 88

RESULT 47
US-09-925-065A-163238
; Sequence 163238, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
```

```
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 163238
; LENGTH: 528
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-163238

Query Match      81.6%; Score 20.4; DB 4; Length 528;
Best Local Similarity 95.5%; Pred. No. 29;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAAAAAAAATCAGCGCCAGGCA 22
   |||||
Db 280 AAAAAAAAAATCAGCGCCAGGCA 301

RESULT 48
US-10-027-632-201503/c
; Sequence 201503, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: Polymorphisms in the Human Genome
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 201503
; LENGTH: 546
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-201503

Query Match      81.6%; Score 20.4; DB 5; Length 546;
Best Local Similarity 95.5%; Pred. No. 29;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAAAAAAAATCAGCGCCAGGCA 22
   |||||
Db 23 AAAAAAAAAATCAGCGCCAGGCA 2

RESULT 49
```


GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 3, 2006, 11:35:31 ; Search time 343.111 Seconds
(without alignments)
61.064 Million cell updates/sec

Title: US-10-719-900-1

Perfect score: 25

Sequence: 1 aaaaaaaatcacggccagcgatg 25

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 6068529 seqs, 419036697 residues

Total number of hits satisfying chosen parameters: 12137058

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database :

Published Applications NA New:
1: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:
2: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:
3: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:
4: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:
5: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:
6: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:
7: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:
8: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:
9: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:
10: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:
11: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	23.4	93.6	201	7	US-10-995-561-79948
C 2	23.4	93.6	41309	7	US-10-995-561-13487
C 3	20.4	81.6	201	7	US-10-995-561-72586
C 4	20.4	81.6	201	7	US-10-995-561-84366
C 5	20.4	81.6	15804	7	US-10-995-561-13294
C 6	20.4	81.6	24446	7	US-10-995-561-13436
C 7	20.4	81.6	40000	7	US-10-995-561-13513
C 8	20.4	81.6	153142	8	US-11-121-086-27
C 9	20.2	80.8	201	7	US-10-995-561-33166
C 10	20.2	80.8	201	7	US-10-995-561-50016
C 11	20.2	80.8	32127	7	US-10-995-561-13352
C 12	20.2	80.8	44229	8	US-11-124-368A-2910
C 13	20.2	80.8	60844	8	US-10-995-561-13359
C 14	20.2	80.8	66280	8	US-11-124-368A-2911
C 15	20.2	80.8	86950	7	US-10-857-780-5
C 16	20.2	80.8	101046	7	US-10-995-561-13330
C 17	20.2	80.8	128978	7	US-10-775-169-345
C 18	20.2	80.8	141121	7	US-10-995-561-13262
C 19	20.2	80.8	160226	8	US-11-121-086-29
C 20	20.2	80.8	168656	8	US-11-112-908-59
C 21	20.2	80.8	168656	8	US-11-112-908-59
C 22	20.2	80.8	170285	8	US-11-112-908-58

C 23	20.2	80.8	179666	8	US-11-121-086-67	Sequence 67, Appl
C 24	20.2	80.8	179892	8	US-11-112-908-39	Sequence 39, Appl
C 25	20.2	80.8	415117	7	US-10-995-561-13274	Sequence 13274, A
C 26	20.2	80.8	415117	7	US-10-995-561-13274	Sequence 13274, A
C 27	19.8	79.2	201	7	US-10-995-561-36951	Sequence 36951, A
C 28	19.8	79.2	201	7	US-10-995-561-50172	Sequence 50172, A
C 29	19.8	79.2	170995	8	US-11-121-086-35	Sequence 35, Appl
C 30	19.8	79.2	201990	7	US-10-995-561-13303	Sequence 13303, A
C 31	19.2	76.8	25	8	US-11-121-849-249080	Sequence 249080, A
C 32	19.2	76.8	25	8	US-11-121-849-250266	Sequence 250266, A
C 33	19.2	76.8	25	8	US-11-121-849-250267	Sequence 250267, A
C 34	19.2	76.8	201	7	US-10-995-561-28311	Sequence 28311, A
C 35	19.2	76.8	201	7	US-10-995-561-28314	Sequence 28314, A
C 36	19.2	76.8	201	7	US-10-995-561-28356	Sequence 28356, A
C 37	19.2	76.8	201	7	US-10-995-561-32893	Sequence 32893, A
C 38	19.2	76.8	201	7	US-10-995-561-32893	Sequence 32893, A
C 39	19.2	76.8	201	7	US-10-995-561-61135	Sequence 61135, A
C 40	19.2	76.8	201	7	US-10-995-561-61137	Sequence 61137, A
C 41	19.2	76.8	201	7	US-10-995-561-61168	Sequence 61168, A
C 42	19.2	76.8	88116	7	US-10-995-561-13351	Sequence 13351, A
C 43	19.2	76.8	222094	7	US-10-995-561-13244	Sequence 13244, A
C 44	19	76.0	201	7	US-10-995-561-44307	Sequence 44307, A
C 45	19	76.0	1125000	7	US-10-995-561-13286	Sequence 13286, A
C 46	18.6	74.4	25	8	US-11-121-849-663017	Sequence 663017, A
C 47	18.6	74.4	201	7	US-10-995-561-35734	Sequence 35734, A
C 48	18.6	74.4	201	7	US-10-995-561-52602	Sequence 52602, A
C 49	18.6	74.4	201	7	US-10-995-561-52678	Sequence 52678, A
C 50	18.6	74.4	201	7	US-10-995-561-78552	Sequence 78552, A
C 51	18.6	74.4	201	8	US-11-124-368A-4303	Sequence 4303, Ap
C 52	18.6	74.4	201	8	US-11-124-368A-4321	Sequence 4321, Ap
C 53	18.6	74.4	16433	8	US-11-124-368A-2875	Sequence 2875, Ap
C 54	18.6	74.4	22855	8	US-11-124-368A-2308	Sequence 2308, Ap
C 55	18.6	74.4	28553	7	US-10-995-561-13477	Sequence 13477, A
C 56	18.6	74.4	35962	7	US-10-775-169-234	Sequence 234, App
C 57	18.6	74.4	38023	7	US-10-995-561-13251	Sequence 13251, A
C 58	18.6	74.4	95832	8	US-11-124-368A-2307	Sequence 2307, A
C 59	18.6	74.4	95832	7	US-10-995-561-13273	Sequence 13273, A
C 60	18.6	74.4	110608	8	US-11-121-086-12	Sequence 12, Appl
C 61	18.6	74.4	119160	8	US-11-121-086-24	Sequence 24, Appl
C 62	18.6	74.4	120096	8	US-11-121-086-49	Sequence 49, Appl
C 63	18.6	74.4	130472	7	US-10-995-561-13312	Sequence 13312, A
C 64	18.6	74.4	149419	8	US-11-112-908-49	Sequence 49, Appl
C 65	18.6	74.4	150038	8	US-11-121-086-23	Sequence 23, Appl
C 66	18.6	74.4	153376	8	US-11-121-086-5	Sequence 5, Appl
C 67	18.6	74.4	156260	8	US-11-121-086-87	Sequence 87, Appl
C 68	18.6	74.4	159146	8	US-11-121-086-49	Sequence 49, Appl
C 69	18.6	74.4	160170	8	US-11-121-086-32	Sequence 32, Appl
C 70	18.6	74.4	161726	8	US-11-112-908-48	Sequence 48, Appl
C 71	18.6	74.4	161726	8	US-11-112-908-52	Sequence 52, Appl
C 72	18.6	74.4	161874	8	US-11-121-086-75	Sequence 75, Appl
C 73	18.6	74.4	166111	8	US-11-112-908-47	Sequence 47, Appl
C 74	18.6	74.4	168516	8	US-11-121-086-3	Sequence 3, Appl
C 75	18.6	74.4	168656	8	US-11-112-908-59	Sequence 59, Appl
C 76	18.6	74.4	170285	8	US-11-112-908-58	Sequence 58, Appl
C 77	18.6	74.4	172147	8	US-11-112-908-22	Sequence 22, Appl
C 78	18.6	74.4	175100	8	US-11-121-086-21	Sequence 21, Appl
C 79	18.6	74.4	176503	8	US-11-121-086-53	Sequence 53, Appl
C 80	18.6	74.4	179777	8	US-11-121-086-106	Sequence 106, App
C 81	18.6	74.4	181172	8	US-11-121-086-41	Sequence 41, Appl
C 82	18.6	74.4	182303	8	US-11-121-086-45	Sequence 45, Appl
C 83	18.6	74.4	189252	8	US-11-121-086-54	Sequence 54, Appl
C 84	18.6	74.4	191684	8	US-11-121-086-2	Sequence 2, Appl
C 85	18.6	74.4	191684	8	US-11-121-086-2	Sequence 2, Appl
C 86	18.6	74.4	197086	8	US-11-121-086-107	Sequence 107, App
C 87	18.6	74.4	197861	8	US-11-112-908-34	Sequence 34, Appl
C 88	18.6	74.4	198161	7	US-10-775-169-52	Sequence 52, Appl
C 89	18.6	74.4	203467	8	US-11-121-086-50	Sequence 50, Appl
C 90	18.6	74.4	212805	8	US-11-112-908-19	Sequence 19, Appl
C 91	18.6	74.4	220895	7	US-10-775-169-88	Sequence 88, Appl
C 92	18.6	74.4	220895	7	US-10-775-169-88	Sequence 88, Appl
C 93	18.6	74.4	246960	8	US-11-121-086-8	Sequence 8, Appl
C 94	18.6	74.4	645179	7	US-10-995-561-13293	Sequence 13293, A
C 95	18.6	74.4	1080000	7	US-10-928-446A-1	Sequence 1, Appl

```

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 79948
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-79948

Query Match      93.6%; Score 23.4; DB 7; Length 201;
Best Local Similarity 96.0%; Pred. No. 0.14;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0

QY 1 AAAAAAAAAATCAGCCGAGCATGG 25
   ||||| ||||| ||||| ||||| |||||
Db 65 AAAAAAAAAATCAGCCGAGCATGG 41

RESULT 2
US-10-995-561-13487/c
; Sequence 13487, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13487
; LENGTH: 41309
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-13487

Query Match      93.6%; Score 23.4; DB 7; Length 41309;
Best Local Similarity 96.0%; Pred. No. 0.44;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0

QY 1 AAAAAAAAAATCAGCCGAGCATGG 25
   ||||| ||||| ||||| ||||| |||||
Db 19143 AAAAAAAAAATCAGCCGAGCATGG 19119

RESULT 3
US-10-995-561-72586/c
; Sequence 72586, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 72586
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-72586

Query Match      81.6%; Score 20.4; DB 7; Length 201;
Best Local Similarity 95.5%; Pred. No. 3.3;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0

QY 1 AAAAAAAAAATCAGCCGAGGCA 22
   ||||| ||||| ||||| ||||| |||||
Db 109 AAAAAAAAAATCAGCCGAGGCA 88

```

```
RESULT 4
US-10-995-561-84366/c
; Sequence 84366, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 84366
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-84366

Query Match      81.6%; Score 20.4; DB 7; Length 201;
Best Local Similarity 95.5%; Pred. No. 3.3;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AAAAAAAAAATCAGGCCAGGCA 22
Db 109 AAAAAAAAAATCAGGCCAGGCA 88

RESULT 5
US-10-995-561-13294/c
; Sequence 13294, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13294
; LENGTH: 15804
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc_feature
; LOCATION: (1)...(15804)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-13294)
US-10-995-561-13294

Query Match      81.6%; Score 20.4; DB 7; Length 15804;
Best Local Similarity 95.5%; Pred. No. 8.3;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AAAAAAAAAATCAGGCCAGGCA 22
Db 58 AAAAAAAAAATCAGGCCAGGCA 37

RESULT 6
US-10-995-561-13436/c
; Sequence 13436, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
```

```
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13436
; LENGTH: 24446
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc_feature
; LOCATION: (1)...(24446)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-13436)
US-10-995-561-13436

Query Match      81.6%; Score 20.4; DB 7; Length 24446;
Best Local Similarity 95.5%; Pred. No. 9.1;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AAAAAAAAAATCAGGCCAGGCA 22
Db 14406 AAAAAAAAAATCAGGCCAGGCA 14385

RESULT 7
US-10-995-561-13513/c
; Sequence 13513, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13513
; LENGTH: 40000
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc_feature
; LOCATION: (1)...(40000)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-13513)
US-10-995-561-13513

Query Match      81.6%; Score 20.4; DB 7; Length 40000;
Best Local Similarity 95.5%; Pred. No. 10;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AAAAAAAAAATCAGGCCAGGCA 22
Db 14069 AAAAAAAAAATCAGGCCAGGCA 14048

RESULT 8
US-11-121-086-27/c
; Sequence 27, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138.60000-00000
; CURRENT APPLICATION NUMBER: US/11/121,086
; CURRENT FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 27
; LENGTH: 153142
```

```
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-27

Query Match      81.6%; Score 20.4; DB 8; Length 153142;
Best Local Similarity 95.5%; Pred. No. 13;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAAAAAAAATCAGCGCCAGGCA 22
   |||||
Db 63127 AAAAAAAAAATCAGCGCCAGGCA 63106

RESULT 9
US-10-995-561-33166/c
; Sequence 33166, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 33166
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-33166

Query Match      80.8%; Score 20.2; DB 7; Length 201;
Best Local Similarity 88.0%; Pred. No. 4.1;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAAAAAATCAGCGCCAGGCA 25
   |||||
Db 76 AAAAAAAAAATCAGCGCCAGGCA 52

RESULT 12
US-10-995-561-13352
; Sequence 13352, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13352
; LENGTH: 32157
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-13352

Query Match      80.8%; Score 20.2; DB 7; Length 32157;
Best Local Similarity 88.0%; Pred. No. 12;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAAAAAATCAGCGCCAGGCA 25
   |||||
Db 12697 AAAAAAAAAATGTCGCCAGGCA 12721

RESULT 13
US-11-124-368A-2910
; Sequence 2910, Application US/11124368A
; Publication No. US20050287559A1
; GENERAL INFORMATION:
; APPLICANT: Michele Cargill
; APPLICANT: James J. Devlin
; APPLICANT: May Luke
; TITLE OF INVENTION: Genetic Polymorphisms Associated with
; TITLE OF INVENTION: Vascular Diseases, Methods of Detection and Uses Thereof
; FILE REFERENCE: CL001524
; CURRENT APPLICATION NUMBER: US/11/124,368A
; CURRENT FILING DATE: 2005-05-09
; PRIOR APPLICATION NUMBER: US 60/568,845
; PRIOR FILING DATE: 2004-05-07
; PRIOR APPLICATION NUMBER: US 60/625,936
; PRIOR FILING DATE: 2004-11-09
```

```
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-27

Query Match      81.6%; Score 20.4; DB 8; Length 153142;
Best Local Similarity 95.5%; Pred. No. 13;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAAAAAAAATCAGCGCCAGGCA 22
   |||||
Db 63127 AAAAAAAAAATCAGCGCCAGGCA 63106

RESULT 9
US-10-995-561-33166/c
; Sequence 33166, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 33166
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-33166

Query Match      80.8%; Score 20.2; DB 7; Length 201;
Best Local Similarity 88.0%; Pred. No. 4.1;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAAAAAATCAGCGCCAGGCA 25
   |||||
Db 201 AAAAAAAAAATTAAGCGCCAGGCA 177

RESULT 10
US-10-995-561-50016/c
; Sequence 50016, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 50016
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-50016

Query Match      80.8%; Score 20.2; DB 7; Length 201;
Best Local Similarity 88.0%; Pred. No. 4.1;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAAAAAATCAGCGCCAGGCA 25
   |||||
Db 176 AAAAAAAAAATTAAGCGCCAGGCA 152

RESULT 11
US-10-995-561-56961/c
```


;; PRIOR FILING DATE: 2004-12-07
;; NUMBER OF SEQ ID NOS: 511
;; SOFTWARE: PatentIn version 3.3
;; SEQ ID NO 59
;; LENGTH: 168656
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-11-112-908-59

Query Match 80.8%; Score 20.2; DB 8; Length 168656;
Best Local Similarity 88.0%; Pred. No. 17;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AAAAAAAAAATCAGCGCCAGGCATGG 25
Db 159496 AAAAAAAAAATCAGCGCCAGGCATGG 159520

RESULT 22

US-11-112-908-58
;; Sequence 58, Application US/11112908
;; Publication No. US20050260659A1
;; GENERAL INFORMATION:
;; APPLICANT: Harris, Cole
;; TITLE OF INVENTION: Breast Cancer Biomarkers
;; FILE REFERENCE: 04-164-US
;; CURRENT APPLICATION NUMBER: US/11/112,908
;; CURRENT FILING DATE: 2005-04-22
;; PRIOR APPLICATION NUMBER: US 60/564,758
;; PRIOR FILING DATE: 2004-04-23
;; PRIOR APPLICATION NUMBER: US 60/575,978
;; PRIOR FILING DATE: 2004-06-01
;; PRIOR APPLICATION NUMBER: US 60/631,702
;; PRIOR FILING DATE: 2004-11-30
;; PRIOR APPLICATION NUMBER: US 60/633,826
;; PRIOR FILING DATE: 2004-12-07
;; NUMBER OF SEQ ID NOS: 511
;; SOFTWARE: PatentIn version 3.3
;; SEQ ID NO 58
;; LENGTH: 170285
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-11-112-908-58

Query Match 80.8%; Score 20.2; DB 8; Length 170285;
Best Local Similarity 88.0%; Pred. No. 17;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AAAAAAAAAATCAGCGCCAGGCATGG 25
Db 159501 AAAAAAAAAATCAGCGCCAGGCATGG 159525

RESULT 23

US-11-121-086-67/c
;; Sequence 67, Application US/11121086
;; Publication No. US20050266459A1
;; GENERAL INFORMATION:
;; APPLICANT: POULSEN, TIM S.
;; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
;; FILE REFERENCE: 09138.6000-00000
;; CURRENT APPLICATION NUMBER: US/11/121,086
;; CURRENT FILING DATE: 2005-05-04
;; PRIOR APPLICATION NUMBER: 60/567,570
;; PRIOR FILING DATE: 2004-05-04
;; NUMBER OF SEQ ID NOS: 107
;; SOFTWARE: PatentIn version 3.3
;; SEQ ID NO 67
;; LENGTH: 179666
;; TYPE: DNA
;; ORGANISM: Homo sapiens

US-11-121-086-67

Query Match 80.8%; Score 20.2; DB 8; Length 179666;
Best Local Similarity 88.0%; Pred. No. 17;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AAAAAAAAAATCAGCGCCAGGCATGG 25
Db 168855 AAAAAAAAAAAGCGCCAGGCATGG 168831

RESULT 24

US-11-112-908-39/c
;; Sequence 39, Application US/11112908
;; Publication No. US20050260659A1
;; GENERAL INFORMATION:
;; APPLICANT: Harris, Cole
;; TITLE OF INVENTION: Breast Cancer Biomarkers
;; FILE REFERENCE: 04-164-US
;; CURRENT APPLICATION NUMBER: US/11/112,908
;; CURRENT FILING DATE: 2005-04-22
;; PRIOR APPLICATION NUMBER: US 60/564,758
;; PRIOR FILING DATE: 2004-04-23
;; PRIOR APPLICATION NUMBER: US 60/575,978
;; PRIOR FILING DATE: 2004-06-01
;; PRIOR APPLICATION NUMBER: US 60/631,702
;; PRIOR FILING DATE: 2004-11-30
;; PRIOR APPLICATION NUMBER: US 60/633,826
;; PRIOR FILING DATE: 2004-12-07
;; NUMBER OF SEQ ID NOS: 511
;; SOFTWARE: PatentIn version 3.3
;; SEQ ID NO 39
;; LENGTH: 179892
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-11-112-908-39

Query Match 80.8%; Score 20.2; DB 8; Length 179892;
Best Local Similarity 88.0%; Pred. No. 17;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AAAAAAAAAATCAGCGCCAGGCATGG 25
Db 101870 AAAAAAAAAAAGCGCGGCATGG 101846

RESULT 25

US-10-995-561-13274
;; Sequence 13274, Application US/10995561
;; Publication No. US20050272054A1
;; GENERAL INFORMATION:
;; APPLICANT: CARGILL, Michele et al.
;; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
;; TITLE OF INVENTION: DETECTION AND USES THEREOF
;; FILE REFERENCE: CL001559
;; CURRENT APPLICATION NUMBER: US/10/995,561
;; CURRENT FILING DATE: 2004-11-24
;; NUMBER OF SEQ ID NOS: 85702
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 13274
;; LENGTH: 415117
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: (1)-(415117)
;; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1
US-10-995-561-13274

Query Match 80.8%; Score 20.2; DB 7; Length 415117;
Best Local Similarity 88.0%; Pred. No. 20;

```
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 AAAAAAAAAATCAGCGCCAGGCATGG 25
Db 268181 AAAAAAAAAAAAAAGCCAGGCATGG 268205

RESULT 26
US-10-995-561-13274/c
; Sequence 13274, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; FILE REFERENCE: CL001559
; CURRENT FILING DATE: 2004-11-24
; CURRENT APPLICATION NUMBER: US/10/995,561
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13274
; LENGTH: 415117
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(415117)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-995-561-13274

Query Match 80.8%; Score 20.2; DB 7; Length 415117;
Best Local Similarity 88.0%; Pred. No. 20;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 AAAAAAAAAATCAGCGCCAGGCATGG 25
Db 289065 AAAAAAAAAAAAAAGCCAGGCATGG 289041

RESULT 27
US-10-995-561-36951/c
; Sequence 36951, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 36951
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-36951

Query Match 79.2%; Score 19.8; DB 7; Length 201;
Best Local Similarity 84.0%; Pred. No. 6.2;
Matches 21; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 1 AAAAAAAAAATCAGCGCCAGGCATGG 25
Db 104 AAAAAAAAAAAAAAGCCAGGCATGG 80

RESULT 28
US-10-995-561-50172/c
; Sequence 50172, Application US/10995561
; Publication No. US20050272054A1
```

```
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 50172
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-50172

Query Match 79.2%; Score 19.8; DB 7; Length 201;
Best Local Similarity 84.0%; Pred. No. 6.2;
Matches 21; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 1 AAAAAAAAAATCAGCGCCAGGCATGG 25
Db 122 AAAAAAAAAAAAAATGCCAGGCRTGG 98

RESULT 29
US-11-121-086-35
; Sequence 35, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138-6000-00000
; CURRENT APPLICATION NUMBER: US/11/121,086
; CURRENT FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 35
; LENGTH: 170995
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-35

Query Match 79.2%; Score 19.8; DB 8; Length 170995;
Best Local Similarity 91.3%; Pred. No. 26;
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 AAAAAAAAAATCAGCGCCAGGCAT 23
Db 111971 AAAAAAAAAAGTCATGCCAGGCAT 111993

RESULT 30
US-10-995-561-13303/c
; Sequence 13303, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13303
; LENGTH: 201990
; TYPE: DNA
; ORGANISM: Homo sapiens
```


Df

CURRENT FILING DATE: 2004-11-24
NUMBER OF SEQ ID NOS: 85702
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 28311
LENGTH: 201
TYPE: DNA
ORGANISM: Homo sapiens
US-10-995-561-28311

Query Match 76.8%; Score 19.2; DB 7; Length 201;
Best Local Similarity 87.5%; Pred.No.12;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AAAAAAAAAATCAGGCCGCGCATG 24
 |||||
Db 134 AAAAAAAAAAAGCCGACGCATG 157
 |||||

RESULT 36
US-10-995-561-28314
Sequence 28314, Application US/10995561
Publication No. US20050272054AI
GENERAL INFORMATION:
APPLICANT: CARGILL, Michele et al.
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
DETECTION AND USES THEREOF
FILE REFERENCE: CL001559
CURRENT APPLICATION NUMBER: US/10/995,561
CURRENT FILING DATE: 2004-11-24
NUMBER OF SEQ ID NOS: 85702
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 28314
LENGTH: 201
TYPE: DNA
ORGANISM: Homo sapiens
US-10-995-561-28314

Query Match 76.8%; Score 19.2; DB 7; Length 201;
Best Local Similarity 87.5%; Pred.No.12;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AAAAAAAAAATCAGGCCGCGCATG 24
 |||||
Db 60 AAAAAAAAAAAGCCGACGCATG 83
 |||||

RESULT 37
US-10-995-561-28356
Sequence 28356, Application US/10995561
Publication No. US20050272054AI
GENERAL INFORMATION:
APPLICANT: CARGILL, Michele et al.
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
DETECTION AND USES THEREOF
FILE REFERENCE: CL001559
CURRENT APPLICATION NUMBER: US/10/995,561
CURRENT FILING DATE: 2004-11-24
NUMBER OF SEQ ID NOS: 85702
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 28356
LENGTH: 201
TYPE: DNA
ORGANISM: Homo sapiens
US-10-995-561-28356

Query Match 76.8%; Score 19.2; DB 7; Length 201;
Best Local Similarity 87.5%; Pred.No.12;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AAAAAAAAAATCAGGCCGCGCATG 24
 |||||
Db 60 AAAAAAAAAAAGCCGACGCATG 83
 |||||

RESULT 38
US-10-995-561-32893/c
Sequence 32893, Application US/10995561
Publication No. US20050272054AI
GENERAL INFORMATION:
APPLICANT: CARGILL, Michele et al.
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
DETECTION AND USES THEREOF
FILE REFERENCE: CL001559
CURRENT APPLICATION NUMBER: US/10/995,561
CURRENT FILING DATE: 2004-11-24
NUMBER OF SEQ ID NOS: 85702
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 32893
LENGTH: 201
TYPE: DNA
ORGANISM: Homo sapiens
US-10-995-561-32893

Query Match 76.8%; Score 19.2; DB 7; Length 201;
Best Local Similarity 87.5%; Pred.No.12;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AAAAAAAAAATCAGGCCGCGCATG 24
 |||||
Db 24 ACATAAAATTGAAGCCGACGCATG 1
 |||

RESULT 39
US-10-995-561-61135/c
Sequence 61135, Application US/10995561
Publication No. US20050272054AI
GENERAL INFORMATION:
APPLICANT: CARGILL, Michele et al.
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
DETECTION AND USES THEREOF
FILE REFERENCE: CL001559
CURRENT APPLICATION NUMBER: US/10/995,561
CURRENT FILING DATE: 2004-11-24
NUMBER OF SEQ ID NOS: 85702
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 61135
LENGTH: 201
TYPE: DNA
ORGANISM: Homo sapiens
US-10-995-561-61135

Query Match 76.8%; Score 19.2; DB 7; Length 201;
Best Local Similarity 87.5%; Pred.No.12;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AAAAAAAAAATCAGGCCGCGCATG 24
 |||||
Db 68 AAAAAAAAAAAGCCGACGCATG 45
 |||||

RESULT 40
US-10-995-561-61137/c
Sequence 61137, Application US/10995561
Publication No. US20050272054AI
GENERAL INFORMATION:
APPLICANT: CARGILL, Michele et al.
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
DETECTION AND USES THEREOF
FILE REFERENCE: CL001559
CURRENT APPLICATION NUMBER: US/10/995,561
CURRENT FILING DATE: 2004-11-24

```
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 61137
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-61137

Query Match          76.8%; Score 19.2; DB 7; Length 201;
Best Local Similarity 87.5%; Pred. No. 12;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAAAAAATCAGGCCAGGCATG 24
   ||||| ||||| ||||| |||||
Db 142 AAAAAAAAAAAGGCCAGGCATG 119

RESULT 41
US-10-995-561-61168/c
; Sequence 61168, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 61168
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-61168

Query Match          76.8%; Score 19.2; DB 7; Length 201;
Best Local Similarity 87.5%; Pred. No. 12;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAAAAAATCAGGCCAGGCATG 24
   ||||| ||||| ||||| |||||
Db 113 AAAAAAAAAAAGGCCAGGCATG 90

RESULT 42
US-10-995-561-13351/c
; Sequence 13351, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13351
; LENGTH: 88116
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-13351

Query Match          76.8%; Score 19.2; DB 7; Length 88116;
Best Local Similarity 87.5%; Pred. No. 42;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAAAAAATCAGGCCAGGCATG 24
   ||||| ||||| ||||| |||||
Db 55029 AAAAAAAAAAAGGCCAGGCATG 55006

RESULT 43
US-10-995-561-13244
; Sequence 13244, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13244
; LENGTH: 222094
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-13244

Query Match          76.8%; Score 19.2; DB 7; Length 222094;
Best Local Similarity 87.5%; Pred. No. 51;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAAAAAATCAGGCCAGGCATG 24
   ||||| ||||| ||||| |||||
Db 163769 AAAAAAAAAAAGGCCAGGCATG 163792

RESULT 44
US-10-995-561-44307/c
; Sequence 44307, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 44307
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-44307

Query Match          76.0%; Score 19; DB 7; Length 201;
Best Local Similarity 90.3%; Pred. No. 14;
Matches 19; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 AAAAAATCAGGCCAGGCATGG 25
   ||||| ||||| ||||| |||||
Db 118 AAAAAATCAGGCCAGGCATGG 98

RESULT 45
US-10-995-561-13286/c
; Sequence 13286, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
```


Mon Feb 6 12:23:11 2006

```
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 78552
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-78552

Query Match      74.4%; Score 18.6; DB 7; Length 201;
Best Local Similarity 84.0%; Pred. No. 22;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy      1 AAAAAAAAAATCAGGCCAGGCATGG 25
        |||||
Db      73 AAAAAAAAAAATAGCCAGGCATGG 49
```

Search completed: February 3, 2006, 16:19:44
Job time : 362.111 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 3, 2006, 10:53:23 ; Search time 863 Seconds
(without alignments)
1646.682 Million cell updates/sec

Title: US-10-719-900-3

Perfect score: 25

Sequence: 1 aaaaacacaccaccactgagctggg 25

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_in.*

3: gb_env.*

4: gb_ov.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pr.*

9: gb_ro.*

10: gb_sts.*

11: gb_sy.*

12: gb_un.*

13: gb_vi.*

14: gb_htg.*

15: gb_pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	25	100.0	4192	9	BC080307	BC080307 Mus muscu
2	25	100.0	4618	9	BC059216	BC059216 Mus muscu
3	25	100.0	211414	9	AC104885	AC104885 Mus muscu
4	21.8	87.2	100983	9	AL663075	AL663075 Mouse DNA
5	21.8	87.2	169405	9	AL596256	AL596256 Mouse DNA
6	21.8	87.2	182785	8	AC145937	AC145937 Pan trogl
7	21.8	87.2	214993	14	AC055773	AC055773 Mus muscu
8	21.8	87.2	276234	14	AC112380	AC112380 Rattus no
9	20.8	83.2	182955	4	AC149756	AC149756 Bos tauru
10	20.2	80.8	2526	5	BC063914	BC063914 xenopus t
11	20.2	80.8	2588	8	BC075515	BC075515 xenopus t
12	20.2	80.8	114676	8	BX323048	BX323048 Human DNA
13	20.2	80.8	128000	8	AC098690	AC098690 Homo sapi
14	20.2	80.8	128638	8	AC015968	AC015968 Homo sapi
15	20.2	80.8	145569	5	AL935194	AL935194 Zebrafish
16	20.2	80.8	155263	14	AC053517	AC053517 Homo sapi
17	20.2	80.8	159976	14	AC016124	AC016124 Homo sapi
18	20.2	80.8	173256	14	AC073301	AC073301 Homo sapi

AC013557	Homo sapi	14	AC013557	80.8	178806	20.2	19	C
AC104828	Homo sapi	8	AC104828	80.8	181489	20.2	20	C
AC099068	Rattus no	14	AC099068	80.8	243279	20.2	21	C
AC095147	Rattus no	14	AC095147	80.8	256914	20.2	22	C
AC160578	Bos tauru	14	AC160578	80.8	290052	20.2	23	C
CR854901	Danio rer	14	CR854901	79.2	192168	19.8	24	C
BX469934	Zebrafish	5	BX469934	79.2	225586	19.8	25	C
AC098538	Rattus no	14	AC098538	79.2	255707	19.8	26	C
M84354	Gallus gall	5	CHRYPEX	77.6	1288	19.4	27	C
K60778	G.gallus DN	5	CHRYPEX	77.6	1288	19.4	28	C
BX548018	Zebrafish	5	BX548018	77.6	113040	19.4	29	C
AC153985	Mus muscu	14	AC153985	77.6	169194	19.4	30	C
AL953906	Mouse DNA	9	AL953906	77.6	172246	19.4	31	C
AC153907	Mus muscu	9	AC153907	77.6	196286	19.4	32	C
CR847836	Danio rer	14	CR847836	77.6	205309	19.4	33	C
CR545469	Danio rer	14	CR545469	77.6	214525	19.4	34	C
AC105669	Rattus no	14	AC105669	77.6	261112	19.4	35	C
AF302834	Haliotis	2	AF302834	76.8	836	19.2	36	C
AL627225	Human DNA	8	AL627225	76.8	3994	19.2	37	C
AL104180	Leishmani	14	AL104180	76.8	90100	19.2	38	C
AL359978	Homo sapi	14	AL359978	76.8	110000	19.2	39	C
Continuation (2 of		14	AL359978_1	76.8	110000	19.2	40	C
Continuation (3 of		14	BX936368_2	76.8	110000	19.2	41	C
CT005261	Leishmani	14	CT005261_0	76.8	110000	19.2	42	C
Continuation (2 of		14	CT005261_1	76.8	110000	19.2	43	C
AL353806	Human DNA	8	AL353806	76.8	111051	19.2	44	C
AC129011	Leishmani	14	AC129011	76.8	124350	19.2	45	C
AC020922	Homo sapi	8	AC020922	76.8	134793	19.2	46	C
AC154498	Mus muscu	14	AC154498	76.8	138861	19.2	47	C
AC145018	Felis cat	14	AC145018	76.8	141663	19.2	48	C
AC157291	Bos tauru	14	AC157291	76.8	154763	19.2	49	C
AC016868	Homo sapi	8	AC016868	76.8	155691	19.2	50	C
AC154788	Mus muscu	14	AC154788	76.8	166692	19.2	51	C
CR354558	Zebrafish	5	CR354558	76.8	177206	19.2	52	C
AC147399	Canis fam	14	AC147399	76.8	178483	19.2	53	C
AL591807	Human Chr	8	CNS07EQ	76.8	183418	19.2	54	C
AC122209	Mus muscu	9	AC122209	76.8	185883	19.2	55	C
AL611949	Mouse DNA	5	AL611949	76.8	186276	19.2	56	C
BX548158	Zebrafish	5	BX548158	76.8	190936	19.2	57	C
AL603708	Mouse DNA	9	AL603708	76.8	198743	19.2	58	C
AL607387	Mus muscu	14	AL607387	76.8	201331	19.2	59	C
AL845296	Mouse DNA	9	AL845296	76.8	203414	19.2	60	C
CT009761	Mus muscu	14	CT009761	76.8	208369	19.2	61	C
AC162504	Bos tauru	14	AC162504	76.8	217178	19.2	62	C
AC123563	Rattus no	14	AC123563	76.8	219934	19.2	63	C
AC144469	Canis fam	14	AC144469	76.8	223041	19.2	64	C
AL732573	Mus muscu	14	AL732573	76.8	223496	19.2	65	C
AC155156	Bos tauru	14	AC155156	76.8	224690	19.2	66	C
AC154520	Mus muscu	14	AC154520	76.8	231513	19.2	67	C
AC096842	Rattus no	14	AC096842	76.8	232927	19.2	68	C
AC140246	Mus muscu	14	AC140246	76.8	250770	19.2	69	C
AC105710	Rattus no	14	AC105710	76.8	280115	19.2	70	C
AC135235	Mus muscu	14	AC135235	76.8	335638	19.2	71	C
AC131401	Rattus no	14	AC131401	76.8	338015	19.2	72	C
BX537351	Zebrafish	5	BX537351	76.0	161005	19	73	C
BV421538	S229P6239	10	BV421538	75.2	510	18.8	74	C
M37920	Homo sapien	8	HUMAMP001	75.2	673	18.8	75	C
BV474964	G591P6285	10	BV474964	75.2	698	18.8	76	C
BV474965	G591P6285	10	BV474965	75.2	735	18.8	77	C
BV485979	S21S56074	10	BV485979	75.2	736	18.8	78	C
BV474961	G591P6285	10	BV474961	75.2	775	18.8	79	C
BV493818	S217P6626	10	BV493818	75.2	775	18.8	80	C
AC100102	Mus muscu	14	AC100102	75.2	44437	18.8	81	C
AL441923	Human DNA	8	AL441923	75.2	76656	18.8	82	C
AC008590	Homo sapi	14	AC008590	75.2	98692	18.8	83	C
AC113637	Rattus no	10	AC113637_0	75.2	110000	18.8	84	C
Continuation (3 of		14	AC113637_2	75.2	110000	18.8	85	C
AC118389	Rattus no	14	AC118389_0	75.2	110000	18.8	86	C
Continuation (2 of		14	AC118389_1	75.2	110000	18.8	87	C
Continuation (130		15	AP008212_129	75.2	110000	18.8	88	C
AP003619	Oryza sat	15	AP003619	75.2	126757	18.8	89	C
AC149475	Zea mays	14	AC149475	75.2	126765	18.8	90	C
AC151050	Zea mays	14	AC151050	75.2	134692	18.8	91	C


```

CDS
276..1310
/genes="Tmem44"
/codon_start=1
/product="Tmem44 protein"
/protein_id="AAH80307.1"
/db_xref="GI:51330402"
/db_xref="GeneID:224090"
/db_xref="MGI:1924489"
/translations="MNPMTLPVCGSKSKSGQSSRRRRRLHRSVPALALPLSL
GPWALWTAIPKAPVPGPQRLLSGLQENPEVGYLLGAIGASWASRIPPPS
ICGKLSLYHLWTRLSALGLLYASAIVAHDPQPEYLLQATPWLISGLRALDLA
IIFLSCVKSIMRSPFAETAEASADTQALTCAEKEENQEAFTDKNSDWLPT
SLSHCKPLRTMTAISRYMELTIEPAQQAGCATRLPGDGTSTRDAASQSPSPYPIQ
VICARVSSSSSSSVSSINSLEQKYWEALNSEQWDPEDVNLKRGDGLLRVQVRRAS
LSPVDLASSD"

ORIGIN
Query Match 100.0%; Score 25; DB 9; Length 4192;
Best Local Similarity 100.0%; Pred. No. 0.33;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAAACACCACTGAGCTGGG 25
|||||
Db 3769 AAAAAACACCACTGAGCTGGG 3793

RESULT 2
BC059216 4618 bp mRNA linear ROD 01-JUN-2005
LOCUS Mus musculus transmembrane protein 44, mRNA (CDNA clone
DEFINITION IMAGE:6408767).
ACCESSION BC059216
VERSION BC059216.1 GI:37747503
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
1 (bases 1 to 4618)
Klausner,R.D., Collins,F.S., Wagner,L.H., Shennan,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,P.,
Datchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Scapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.D., Hulyk,S.W., R.A.,
Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Kettman,M., Madan,A., Rodrigues,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,M.C., Blakeley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.G., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,F.S., Krzywinski,M.I., Skalska,U., Smalius,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Mammalian Gene Collection Program Team
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 4618)

NIH MGC Project
Direct Submission
Submitted (01-OCT-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa

CDNA Library Preparation: M. Bento Soares, University of Iowa
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www.shgc.stanford.edu
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAC Place: 125 Row: 9 Column: 11.
Location/Qualifiers
1..4618
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CS7BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:6408767"
/tissue_type="Brain, enriched mouse brain 12.5dp"
/clone_lib="NIH BMAP_F00"
/lab_host="DH10B"
/note="Vector: pYX-ASC"

FEATURES
source
Query Match 100.0%; Score 25; DB 9; Length 4618;
Best Local Similarity 100.0%; Pred. No. 0.33;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAAACACCACTGAGCTGGG 25
|||||
Db 4180 AAAAAACACCACTGAGCTGGG 4204

RESULT 3
AC104885 211414 bp DNA linear ROD 07-NOV-2003
LOCUS Mus musculus chromosome 16, clone RP23-125P19, complete sequence.
DEFINITION AC104885
ACCESSION AC104885
VERSION AC104885.8 GI:38198721
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Muridae; Murinae; Mus.
1 (bases 1 to 211414)
Birren,B., Nusbaum,C. and Lander,E.
Mus musculus chromosome 16, clone RP23-125P19
Unpublished
2 (bases 1 to 211414)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhgalter,B.,
Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
Choepe,Y., Colangelo,M., Collins,S., Collamore,A., Cook,A.,
Cooke,P., DeAcellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliiev,I., Johnson,R.,
Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K.,
Lamazar,S., Landers,T., Lehotzky,J., Levine,R., Liu,G.,
MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrum,J.,
Meneus,L., Mihova,T., Mienga,V., Murphy,T., Naylor,J., Nguyen,C.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,J., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupack,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigglio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

```



```

repeat_region complement(12158..12455)
repeat_region /rpt_family="RMER19B"
repeat_region complement(12460..13065)
repeat_region /rpt_family="RMER19B"
repeat_region 13066..13254
repeat_region /rpt_family="B2 Mmi"
repeat_region complement(13255..13285)
repeat_region /rpt_family="RMER19B"
repeat_region complement(13520..13652)
repeat_region /rpt_family="RMER15"
repeat_region complement(13736..13895)
repeat_region /rpt_family="RLTR11A"
repeat_region complement(14020..14224)
repeat_region /rpt_family="BGLII"
repeat_region complement(14225..14325)
repeat_region /rpt_family="RMER15"
repeat_region 14712..14733

Query Match 100.0%; Score 25; DB 9; Length 211414;
Best Local Similarity 100.0%; Pred. No. 0.36;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAAACACACCACTGAGCTGG 25
Db 58978 AAAAAACACACCACTGAGCTGG 58954

RESULT 4
AL663075/c
LOCUS Mouse DNA sequence from clone RP23-430121 on chromosome 11,
DEFINITION complete sequence.
ACCESSION AL663075
VERSION AL663075.6 GI:19572106
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 100983)
Sycamore.N.
Direct Submission
Submitted (04-FEB-2005) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vegas@sanger.ac.uk
Clone requests: clonerequest@sanger.ac.uk
On Mar 21, 2002 this sequence version replaced gi:19031807.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em.; EMBL; Sw.; SWISSPROT; Tr.; TREMBL; Wp.; WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep -----
Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: vegas@sanger.ac.uk
-----
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one subclone; and the assembly was confirmed by restriction digest,
except on the rare occasion of the clone being a YAC.
Sequence from the Mouse Genome Sequencing Consortium whole genome
shotgun may have been used to confirm this sequence. Sequence data
from the whole genome shotgun alone has only been used where it has
a phred quality of at least 30.
RP23-430121 is from the RPCI-23 Mouse BAC Library
constructed by the group of Pieter de Jong.
For further details see http://www.chori.org/bacpac/home.htm
VECTOR: pBAC3.6.

```

```

FEATURES
source
Location/Qualifiers
1..100983
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosomes="11"
/clone_lib="RPCI-23"
2761..72904
/locus_tag="RP23-430121.3-001"
/pseudo
2761..72904
/locus_tag="RP23-430121.3-001"
/notes="match: proteins: Sw:Q28894 Sw:Q81UB2"
/pseudo
/codon_start=1
/product="extracellular proteinase inhibitor (Expi)
pseudogene"
5461..7738
/genes="RP23-430121.1"
/locus_tag="RP23-430121.1-001"
join(5461..5643,6264..6412,7538..7738)
/genes="RP23-430121.1"
/locus_tag="RP23-430121.1-001"
/product="novel protein similar to extracellular
proteinase inhibitor Expi"
/notes="match: ESTs: Em:AA201047.1 Em:AA237876.1
Em:AA244943.1 Em:AA274581.1 Em:AA561990.1 Em:AA666504.1
Em:AA981470.1 Em:BE624328.1 Em:CB319028.1 Em:W08116.1
match: cDNAs: Em:BC010986.1 Em:BC038870.1"
join(5562..5643,6264..6412)
/genes="RP23-430121.1"
/locus_tag="RP23-430121.1-001"
/standard_name="OTTMUSP0000000924"
/notes="match: proteins: Sw:P14730 Sw:Q62477 Tr:Q91VQ6
Tr:Q99M36"
/codon_start=1
/product="novel protein similar to extracellular
proteinase inhibitor Expi"
/protein_id="CAI25670.1"
/db_xref="GI:56206511"
/db_xref="InterPro:IPR008197"
/translation="MKTATVFLVALITVGMNTTYVYVSCPKEFKGACPKPSPESVG
ICVDQSGGSGCPGNMKCCNSCGHCKTPVF"
7119..7724
/genes="RP23-430121.1"
/locus_tag="RP23-430121.1-001"
7738
/genes="RP23-430121.1"
/locus_tag="RP23-430121.1-001"
join(10485..10669,11327..11469,12626..12810)
/locus_tag="RP23-430121.2-001"
join(10485..10669,11327..11469,12626..12810)
/locus_tag="RP23-430121.2-001"
/product="extracellular proteinase inhibitor"
/notes="match: ESTs: Em:AW208862.1 Em:BB823505.1
Em:BB828561.1 Em:BB835917.1 Em:BE282859.1 Em:BE553674.1
Em:BE554305.1 Em:BF181080.1 Em:BG084226.2 Em:BG863639.1
Em:BI111742.1 Em:BI453987.1 Em:BI660358.1 Em:CB598051.1
match: cDNAs: Em:AF159701.1 Em:BC02038.1 Em:BC010986.1
Em:X93037.1"
join(10588..10669,11327..11469)
/locus_tag="RP23-430121.2-001"
/standard_name="OTTMUSP0000000922"
/notes="match: proteins: Sw:Q62477 Tr:Q91VQ6 Tr:Q99M36"
/codon_start=1
/product="extracellular proteinase inhibitor"
/protein_id="CAI25671.1"
/db_xref="GI:56206512"
/db_xref="InterPro:IPR008197"
/translation="MKTATVFLVALITVGMNTTAWALSNPKFKGACPKPPPSFGTC
DERCTGDGSGGNMKCCNSCGHCKTPVF"

```

```
gene      join(<10593. .10669,11327. .11469,11569. .11698,12626. .12809)
          /locus_tag="RP23-430I21.2-002"
mRNA      join(<10593. .10669,11327. .11469,11569. .11698,12626. .12809)
          /locus_tag="RP23-430I21.2-002"
          /product="extracellular proteinase inhibitor"
          /note="match: ESTs: Em:BI453077.1"
          join(<10593. .10669,11327. .11469)
          /locus_tag="RP23-430I21.2-002"
          /standard_name="OTTMUSP0000000923"
          /codon_start=2
          /product="extracellular proteinase inhibitor"
          /protein_id="CAI25672.1"
          /db_xref="GI:56206513"
          /db_xref="InterPro:IPR002308"
          /db_xref="InterPro:IPR008197"
          /db_xref="UniProt/TREMBL:Q5SS19"
          /translation="TATVFLVLIIFMTWTAWALSNPKKPGACPKPPRPFSGTCDE
          RCTGDCSGNNKCCSGNGCHACKPPVF"
          join(10595. .10669,11327. .12168)
          /locus_tag="RP23-430I21.2-003"
          join(10595. .10669,11327. .12168)
          /locus_tag="RP23-430I21.2-003"
          /product="extracellular proteinase inhibitor"
          /note="match: ESTs: Em:BF137096.1 Em:BF139756.1
          Em:BG863008.1"
          12790. .12795
          /locus_tag="RP23-430I21.2-002"
          12810
          /locus_tag="RP23-430I21.2-001"
          48410. .54112
          /locus_tag="RP23-430I21.5-001"
          join(48410. .48593,50703. .50803,53964. .54112)
          /locus_tag="RP23-430I21.5-001"
          /product="novel protein"
          /note="match: ESTs: Em:AV006463.1 Em:AW109625.1
          Em:BG633723.1 Em:BY080598.1 Em:BY082582.1 Em:BY085790.1
          Em:BY089336.1 Em:BY094209.1 Em:BY094545.1 Em:BY102260.1
          Em:BY449514.1 Em:BY454341.1 Em:W11247.1
          match: cDNAs: Em:AK003164.1 Em:AK003352.1"
          join(48497. .48593,50703. .50797)
          /locus_tag="RP23-430I21.5-001"
          /standard_name="OTTMUSP0000000928"
          /note="match: proteins: Tr:QB8TE6"
          /codon_start=1
          /product="novel protein"
          /protein_id="CAI25673.1"
          /db_xref="GI:56206514"
          /db_xref="InterPro:IPR008197"
          /db_xref="UniProt/TREMBL:QB8TE6"
          /translation="MKLGAFLLVSLITLSLEVQELQAAVRPLQLLTCAELCRGDWD
          CQPEQCVSIGCSHCTYN"
          54088. .54093
          /locus_tag="RP23-430I21.5-001"
          54108
          /locus_tag="RP23-430I21.5-001"
          join(55166. .55411,57180. .57287,58780. .58920,59759. .59874,
          60774. .60888,62369. .62470,66356. .66493,67323. .67630,
          68548. .68725,70526. .70732,71138. .71237,73715. .74070,
          75340. .75462,75892. .75978,77171. .77305,78084. .78206,
          78708. .78832,80315. .80411,80871. .81075,82621. .85224)
          /locus_tag="RP23-430I21.6-001"
          join(55166. .55411,57180. .57287,58780. .58920,59759. .59874,
          60774. .60888,62369. .62470,66356. .66493,67323. .67630,
          68548. .68725,70526. .70732,71138. .71237,73715. .74070,
          75340. .75462,75892. .75978,77171. .77305,78084. .78206,
          78708. .78832,80315. .80411,80871. .81075,82621. .85224)
          /locus_tag="RP23-430I21.6-001"
          /product="novel protein"
          /note="match: ESTs: Em:AK638885.1 Em:AI747047.1
          Em:AM646278.1 Em:BB613238.1 Em:BE310138.1 Em:BEF143860.1
          Em:BI688043.1 Em:BI693858.1 Em:BQ561083.1 Em:BU919386.1
          Em:BY010527.1 Em:BY711395.1 Em:BY723820.1 Em:BY732042.1
          Em:BY741972.1 Em:CB525352.1 Em:CD350433.1 Em:CD545893.1
          Em:BY741972.1 Em:CB525352.1 Em:CD350433.1 Em:CD545893.1
```

Em:CF162117.1 Em:CF170063.1 Em:CF733576.1
match: cDNAs: Em:AK012216.1 Em:CG784573.1"
join(55184. .55615,57180. .57287,58780. .58920,59759. .59874,
60774. .60888,62369. .62470,66356. .66493,67323. .67630,
68548. .68725,70526. .70732,71138. .71237,73715. .74070,
75340. .75462,75892. .75978,77171. .77305,78084. .78206,
78708. .78832,80315. .80411,80871. .81075,82621. .83491)
/locus_tag="RP23-430I21.6-002"
join(55184. .55615,57180. .57287,58780. .58920,59759. .59874,
60774. .60888,62369. .62470,66356. .66493,67323. .67630,
68548. .68725,70526. .70732,71138. .71237,73715. .74070,
75340. .75462,75892. .75978,77171. .77305,78084. .78206,
78708. .78832,80315. .80411,80871. .81075,82621. .83491)
/locus_tag="RP23-430I21.6-002"

Query Match 87.2%; Score 21.8; DB 9; Length 100983;
Best Local Similarity 92.0%; Pred. No. 14;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAACACCACTGAGCTGGG 25
|||||
Db 22931 AAAAAAATCACCACGCTGGG 22907
|||||

RESULT 5
AL596256
LOCUS
DEFINITION
complete sequence.
AL596256 169405 bp DNA linear ROD 09-FEB-2005
Mouse DNA sequence from clone RP23-259D24 on chromosome 11,
complete sequence.
AL596256 GI:18855209
VERSION
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 169405)
Brown, J.
Direct Submission
Submitted (04-FEB-2005) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vegas@sanger.ac.uk
Clone requests: clonerequest@sanger.ac.uk
On Feb 21, 2002 this sequence version replaced gi:18135107.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep -----
Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: vegas@sanger.ac.uk

This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one subclone; and the assembly was confirmed by restriction digest,
except on the rare occasion of the clone being a YAC.
Sequence from the Mouse Genome Sequencing Consortium whole genome
shotgun may have been used to confirm this sequence. Sequence data
from the whole genome shotgun alone has only been used where it has
a phred quality of at least 30.
RP23-259D24 is from the RPCI-23 Mouse BAC Library
constructed by the group of Pieter de Jong.
For further details see http://www.chori.org/bacpac/home.htm
VECTOR: pBACes3.6.
Location/Qualifiers
1. 169405
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="11"
/clone="RP23-259D24"

```
gene
/c/one lib="RPC1-23"
join(complement(44425..44567),complement(42239..42473),
complement(32825..32993),complement(18848..18993),
complement(17953..18150),complement(16133..16327),
complement(15412..15574),complement(14767..14877),
complement(6313..6966),AL592065.7:6124..6337,
AL592065.7:7974..8052,AL592065.7:11381..11502,
AL592065.7:13481..13571,AL592065.7:13692..13906,
AL592065.7:14190..14386,AL592065.7:15840..16756,
AL592065.7:21832..23993,AL592065.7:25880..26103,
AL592065.7:26756..26945,AL592065.7:28146..28602,
AL592065.7:29704..29923,AL592065.7:31094..31285,
AL592065.7:31613..31836,AL592065.7:35301..35443,
AL592065.7:36114..36272,AL592065.7:36381..36566,
AL592065.7:38068..38216,AL592065.7:39489..39662,
AL592065.7:42473..42573,AL592065.7:43918..48003)
/locus_tag="RP23-5084.3-001"
join(complement(44425..44567),complement(42239..42473),
complement(32825..32993),complement(18848..18993),
complement(17953..18150),complement(16133..16327),
complement(15412..15574),complement(14767..14877),
complement(6313..6966),AL592065.7:6124..6337,
AL592065.7:7974..8052,AL592065.7:11381..11502,
AL592065.7:13481..13571,AL592065.7:13692..13906,
AL592065.7:14190..14386,AL592065.7:15840..16756,
AL592065.7:21832..23993,AL592065.7:25880..26103,
AL592065.7:26756..26945,AL592065.7:28146..28602,
AL592065.7:29704..29923,AL592065.7:31094..31285,
AL592065.7:31613..31836,AL592065.7:35301..35443,
AL592065.7:36114..36272,AL592065.7:36381..36566,
AL592065.7:38068..38216,AL592065.7:39489..39662,
AL592065.7:42473..42573,AL592065.7:43918..48003)
/locus_tag="RP23-5084.3-001"
/product="novel protein"
/note="match: ESTs: Em:BB152232.2 Em:BB329744.1
Em:BB630436.1 Em:BG091148.1 Em:BU924606.1 Em:CA532211.1
Em:CA890306.1 Em:CD349776.1 Em:CD551607.1 Em:CF163861.1
Em:CF741768.1 Em:CF744131.1 Em:CF892872.1 Em:CF906293.1
match: CDNAs: Em:AB011165.2 Em:AF117754.1"
join(complement(44425..44900),complement(42239..42473),
complement(32825..32993),complement(18848..18993),
complement(17953..18150),complement(16133..16327),
complement(15412..15574),complement(14767..14877),
complement(6313..6966),AL592065.7:6124..6337,
AL592065.7:7974..8052,AL592065.7:11381..11502,
AL592065.7:13481..13571,AL592065.7:13692..13906,
AL592065.7:14190..14386,AL592065.7:15840..16756,
AL592065.7:21832..23993,AL592065.7:25880..26103,
AL592065.7:26756..26945,AL592065.7:28146..28602,
AL592065.7:29704..29923,AL592065.7:31094..31285,
AL592065.7:31613..31836,AL592065.7:35301..35443,
AL592065.7:36114..36272,AL592065.7:36381..36566,
AL592065.7:38068..38216,AL592065.7:39489..39662,
AL592065.7:42473..42573,AL592065.7:43918..44050)
/locus_tag="RP23-5084.3-001"
/standard_name="OTTMUSP0000001175"
/note="match: proteins: Sw:Q9UHV7 Tr:BA25519 Tr:BAC97978
Tr:Q80TM3 Tr:Q810K4 Tr:Q8C2D6 Tr:Q9H8C0 Tr:Q9NSY9
Tr:Q9UFD8 Tr:Q9UPX5"
/codon_start=1
/product="novel protein"
/db_xref="GI:56205831"
/translataion="MSSFSVNGASLEDCNLCFLADLTGKIKRWVWGQTSAPIL
FPVTEEDILSSFRCLKADVLGVWRDQRPGRRELWIFWGDKNPADIHDLSE
EDWVNGSLYECLTLFKAHNLLECLRNFRKFWKFEYEDKPIKNSHEL
SCSFTFLHNGSVNCTVEINQHPVYLLSEHVTLAQONSFPVITLSPFLNGTLT
QOAFKMSDATKILIGKQFYPI SCGLKEMSEKQDDMDWEDSLAAVEVLVAGVEM
IYPCFVLVPSQDILPAPSSVASHCSACSLGIHQVPASTRDPAMSVTLTPSPPEV
QTVDPQSAQKWEAFSSVSDGSDTSHHGKIKRKLHNVHVWQECNNRQNRK
KYSATSSGLKEETADKIGWDFEATQRTSCSLRHKLSKTRNTGQAGKSGQQQ
QVLPKHTNEKQDKSEKPKRPLTPFHHRVSVSDIEMDITDSQRLVISAADSQVFP
SNIRNDVAKTPQMHTELANSPPPLSPHPCDVDEGVTKTFTSPQHFYQMTPP
```

gene

mRNA

CDS

Query Match 87.2%; Score 21.8; DB 9; Length 169405;
Best Local Similarity 92.0%; Pred. No. 14;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAACACCACTGAGCTGGG 25
|||||||
Db 159735 AAAAAACACCACTGAGCTGGG 159759
|||||||

RESULT 6
AC145937/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AC145937
Pan troglodytes BAC clone RP43-10P16 from chromosome 7, complete
sequence.
AC145937.2 GI:36016669
HTG.
Pan troglodytes (chimpanzee)
Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Pan.
1 (bases 1 to 182785)
Tomlinson, C., Cotton, M. and Kozlowicz, A.
The sequence of Pan troglodytes BAC clone RP43-10P16
Unpublished (2001)
2 (bases 1 to 182785)
Wilson, R.K.
Direct Submission
Submitted (01-AUG-2003) Genetics, Genome Sequencing Center, 4444
Forest Park Parkway, St. Louis, MO 63108, USA
3 (bases 1 to 182785)
Wilson, R.
Direct Submission
Submitted (26-SEP-2003) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
4 (bases 1 to 182785)

AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Wilson, R.K.
 Direct Submission
 Submitted (16-NOV-2004) Genetics, Genome Sequencing Center, 4444
 Forest Park Parkway, St. Louis, MO 63108, USA
 5 (bases 1 to 182785)
 Wilson, R.K.
 Direct Submission
 Submitted (27-APR-2005) Washington University School of Medicine,
 Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO
 63108, USA
 On Sep 26, 2003 this sequence version replaced gi:33386887.
 ----- Genome Center
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: <http://genome.wustl.edu>
 Contact: submissions@wustl.edu
 ----- Summary Statistics
 Center project name: C_PT010P16

NOTICE:

This sequence was finished as follows unless otherwise noted:
 all regions were double stranded, sequenced with an alternate
 chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by sequence
 from more than one subclone; and the assembly was confirmed by
 restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. Wes Warren,
 Department of Genetics, Washington University, St. Louis MO. For
 additional information about the map position of this sequence, see
<http://genome.wustl.edu>

SOURCE INFORMATION:

The RPCI-43 BAC Library has been constructed by Chung-Li Shu. DNA
 was isolated from white blood cells obtained from a male chimpanzee
 (Pan troglodytes, 'Clint', Yerkes #C0471; birthdate: 6-6-80). The
 clone and detailed information can be obtained from ResGen
 (<http://www.resgen.com>) or Pieter de Jong and co-workers at
<http://www.bacpac.chori.org>.

This sequence is the entire insert of the clone.

FEATURES

source
 Location/Qualifiers
 1..182785
 /organism="Pan troglodytes"
 /mol_type="genomic DNA"
 /db_xref="taxon:9598"
 /chromosomes="7"
 /clone_lib="RPCI-43"
 /clone="RP43-10P16"
 14981..15169
 /notes="Unresolved simple sequence repeat."
 26038..26707
 /notes="Unresolved tandem repeat."

ORIGIN

Query Match 87.2%; Score 21.8; DB 8; Length 182785;
 Best Local Similarity 92.0%; Pred. No. 14;
 Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAACACCACTGAGCTGGG 25

Db 137382 AAAAAACACCACTCATCTGGG 137358

RESULT 7

AC055773 214993 bp DNA linear HTG 13-MAR-2003
 LOCUS
 DEFINITION Mus musculus clone RP23-186B19, WORKING DRAFT SEQUENCE, 3 unordered
 pieces.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AC055773.4 GI:28933881
 HTG; HTGS_PHASE1; HTGS_DRAFT.
 Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muridea; Muridae; Murinae; Mus.

REFERENCE

AUTHORS
TITLE
JOURNAL

Birren, B., Nusbaum, C. and Lander, E.
 Mus musculus, clone RP23-186B19

REFERENCE

AUTHORS

2 (bases 1 to 214993)
 Unpublished
 Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,
 Anderson, S., Baldwin, J., Barna, N., Bastien, V., Beda, F.,
 Boguslavsky, L., Boukhgalter, B., Brown, A., Burkett, G.,
 Campiano, A., Castle, A., Choepel, Y., Colangelo, K., Collins, S.,
 Collymore, A., Cooke, P., DeArellano, K., Dewar, K., Diaz, J. S.,
 Dodge, S., Domino, M., Doyle, M., Ferreira, P., FitzHugh, W., Gage, D.,
 Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,
 Grand-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L.,
 Howland, J. C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,
 Klein, J., LaRoque, K., Lamazares, R., Landers, T., Lehoczy, J.,
 Levine, R., Lieu, C., Liu, G., Locke, K., Macdonald, P., Marquis, N.,
 McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPheeters, R.,
 Meldrim, J., Meneus, L., Mihova, T., Miranda, C., Mlenga, V., Morrow, J.,
 Murphy, T., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P.,
 O'Neill, D., Oliver, T. M., Oliver, J., Peterson, K., Pierre, N.,
 Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,
 Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,
 Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
 Testaye, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J.,
 Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J.,
 Young, G., Zainoun, J., Zimmer, A. and Zody, M.

TITLE**JOURNAL**

Submitted (18-APR-2000) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA

3 (bases 1 to 214993)

REFERENCE**AUTHORS**

Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N.,
 Anderson, S., Arachchi, H. M., Barna, N., Bastien, V., Bloom, T.,
 Boguslavsky, L., Boukhgalter, B., Camarata, J., Chang, J., Choepel, Y.,
 Collymore, A., Cook, A., Cooke, P., Corum, B., DeArellano, K.,
 Diaz, J. S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S.,
 Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S.,
 Gerdmann, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B.,
 Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
 Kamat, A., Karatas, A., Keils, C., Landers, T., Levine, R.,
 Lindblad-Toh, K., Liu, G., Lui, A., Mabbitt, R., MacLean, C.,
 Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M.,
 Meldrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J.,
 Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P.,
 O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N.,
 Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P.,
 Roman, J., Schauer, S., Schupback, R., Seaman, S., Severy, P., Smith, C.,
 Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M.,
 Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M.,
 Vassiliev, H., Venkataraman, V. S., Viel, R., Vo, A., Wilson, B., Wu, X.,
 Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

COMMENT

Submitted (13-MAR-2003) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA

On Mar 13, 2003 this sequence version replaced gi:13123353.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIER

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L5031

Center clone name: 186_B_19

```
----- Summary Statistics
Sequencing vector: M13; M77815; 37% of reads
Sequencing vector: Plasmid; n/a; 63% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 214575 bases at least Q40
Consensus quality: 214693 bases at least Q30
Consensus quality: 214775 bases at least Q20
Insert size: 220000; agarose-fp
Insert size: 214793; sum-of-contigs
Quality coverage: 12.1 in Q20 bases; agarose-fp
Quality coverage: 12.4 in Q20.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 67177: contig of 67177 bp in length
* 67178 67277: gap of 100 bp
* 67278 91121: contig of 23844 bp in length
* 91121 91221: gap of 100 bp
* 91222 214993: contig of 123772 bp in length.
FEATURES
source
1. 214993
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="RP23-186B19"
/clone_lib="RPCI-23 Female Mouse BAC"
misc_feature
1..67177
/note="assembly_fragment"
clone_end:SP6
vector_side:left
67178..67277
/estimated_length=100
misc_feature
67278..91121
/note="assembly_fragment"
91122..91221
/estimated_length=100
misc_feature
91222..214993
/note="assembly_fragment"
clone_end:T7
vector_side:right
ORIGIN
Query Match 87.2%; Score 21.8; DB 14; Length 214993;
Best Local Similarity 92.0%; Pred. No. 14;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 AAAAAAACACCACTGAGCTGG 25
||||| ||||| ||||| |||||
Db 65299 AAAAAAAATCACCCTGAGCTGG 65323
AC112380 276234 bp DNA linear HTG 22-SEP-2002
Rattus norvegicus clone CH230-24L19, *** SEQUENCING IN PROGRESS
***, 2 unordered pieces.
AC112380
AC112380.3 GI:23195384
HTG; HTGS PHAS1; HTGS DRAFT; HTGS_ENRICHED.
KEYWORDS
Rattus norvegicus (Norway rat)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Rattus.
1 (bases 1 to 276234)
REFERENCE
Muzny,D,Marie., Metzker,M, Lee., Abramson,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
Balduino,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
Egan,A., Escoto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
Greggorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
Gunaratne,P., Haaland,W., Hamill,C., Hamilton,C., Hamilton,K.,
Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
Karpachy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
Lorensuhewa,L., Louisedge,H., Lozado,R.J., Lu,X., Ma,J.,
Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Parks,K.,
Nwaokemele,O., Okwuonu,G., Olarnpungoon,A., Pal,S., Parks,K.,
Pasternak,S., Paul,H., Perez,A., Perez,L., Prannkoch,C.,
Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
Puzo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
Sanders,W., Savary,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Umami,K.,
Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
Williams,G., Willson,R., Wleczek,R., Woodden,H., Worley,K.,
Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
Weinstock,G. and Gibbs,R.A.
Direct Submission
Unpublished
2 (bases 1 to 276234)
Worley,K.C.
Direct Submission
Submitted (22-FEB-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 276234)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (22-SEP-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Sep 19, 2002 this sequence version replaced gi:21744211.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). As a result, the
sequence may extend beyond the ends of the clone and there may be
contigs that consist entirely of whole genome shotgun sequence
reads. Both end sequences and whole genome shotgun sequence only
contigs will be indicated in the feature table.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
```

```

Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GORW
Center clone name: CH230-24L19
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 211787 bases at least Q40
Consensus quality: 214627 bases at least Q30
Consensus quality: 216603 bases at least Q20
Estimated insert size: 231080; sum-of-contigs estimation
Quality coverage: 4x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
consists of 2 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.
*
* 15842: contig of 15842 bp in length
* 15843 15942: gap of unknown length
* 15943 276234: contig of 260292 bp in length.
Location/Qualifiers
1..276234
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-24L19"
4375..4751
/note="clone_boundary
clone end:Sp6
site:EcoRI
end_sequence:BH281496"
15843..15942
/estimated_length=unknown
15943..20071
/note="wgs contig"
complement(259013..259418)
/note="clone_boundary
clone end:T7
site:EcoRI
end_sequence:BH281495"
ORIGIN
Query Match 87.2%; Score 21.8; DB 14; Length 276234;
Best Local Similarity 92.0%; Pred. No. 15;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AAAAAACACCACTGAGCTGGG 25
||| ||||| ||||| ||||| |||||
Db 5299 AAGAAACACCACTGAGCTGGG 5275
||| ||||| ||||| ||||| |||||

RESULT 9
AC149756
LOCUS
DEFINITION
Bos taurus BAC CH240-36112 (Children's Hospital Oakland Research
Institute Bovine BAC Library (male)) complete sequence.
ACCESSION
AC149756
VERSION
AC149756.4 GI:50872227
KEYWORDS
HTG.
SOURCE
Bos taurus (cow)
ORGANISM
Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
Pecora; Bovidae; Bovinae; Bos.
REFERENCE
1 (bases 1 to 182955)
Muzny,D.Marie., Metzker,M.Lee., Abramson,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
Anyalebechi,V., Ayoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,H., Dederick,D.,
Delgado,M.O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
Gregregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,M.,
Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
Hernandez,R., Hines,S., Hladun,S.I., Hodgson,A., Hogues,M.,
Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
Karpachy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
Lorensuhewa,L., Louissegh,H., Lozado,R.J., Lu,X., Ma,J.,
Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
Mawhinney,S., McLeod,M., McNeill,T., Meenen,E., Milosavljevic,A.,
Miner,G., Minja,E., Montemayor,J., Moore,S., Morgan,M., Morris,K.,
Morris,S., Munidasa,M., Murphy,M., Nair,L., Nankervis,C., Neal,D.,
Newton,N., Nguyen,N., Norris,S., Nwaokemele,O., Okwono,G.,
Olarnpunsagoon,A., Pal,S., Parks,K., Pasternak,S., Paul,H.,
Perez,A., Perez,L., Pfannkoch,C., Plopper,F., Poindexter,A.,
Popovic,D., Primus,E., Pu,L.-L., Puazo,M., Quiroz,J., Rachlin,E.,
Reaves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y.,
Reuter,M., Richards,S., Riggs,F., Rives,C., Rodkey,T., Rojas,A.,
Rose,M., Rose,R., Ruiz,S.J., Sanders,W., Savory,G., Scherer,S.,
Scott,G., Shatsman,S., Shen,H., Shetty,J., Shvartsbeyn,A.,
Sisson,I., Sitter,C.D., Smajs,D., Sneed,A., Sodergren,E.,
Song,X.-Z., Sorelle,R., Sosa,J., Steidle,M., Strong,R., Sutton,A.,
Svatek,A., Tabor,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S.,
Tingey,A., Trejos,Z., Usmani,K., Valas,R., Vera,V., Villasana,D.,
Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J.,
Warren,R., Wei,X., White,F., Williams,G., Willson,R., Wisczyk,R.,
Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S.,
Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,J., Zhou,X.,
Zhao,S., Dunn,D., von Niederhausern,A., Weiss,R., Smith,D.R.,
Holt,R.A., Smith,H.O., Weinstock,G. and Gibbs,R.A.
Direct Submission
Unpublished
2 (bases 1 to 182955)
Worley,K.C.
Direct Submission
Submitted (19-JUN-2004) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 182955)
Worley,K.C.
Direct Submission
Submitted (28-JUL-2004) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
4 (bases 1 to 182955)
Worley,K.C.
Direct Submission
Submitted (31-JUL-2004) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
COMMENT
On Jul 31, 2004 this sequence version replaced gi:50428583.
Sequencing is completed to a minimum standard of double strand
coverage with a minimum of 2 clones and 2 reads with no ambiguities
or 2 chemistries with a minimum of 2 clones and 3 reads with no
ambiguities. If the sequence quality does not meet this standard,
it will be indicated in the annotation.
FEATURES
Location/Qualifiers
1..182955
source

```


Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettman,M., Madan,A., Rodrigues,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Mammalian Gene Collection Program Team
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 2526)
Klein,S. and Gerhard,D.S.
Direct Submission
Submitted (08-DEC-2003) National Institutes of Health, Xenopus Gene Collection (XGC), National Institute of Child Health and Human Development, 6100 Executive Boulevard, Room 4B01, Rockville, MD 20892-7510, USA
NIH-MGC Project
Contact: XGC help desk
Email: cgabbs@mail.nih.gov
Tissue Procurement: Robert M. Grainger
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca
Martin Hirst, Thomas Zeng, Ryan Morin, Michelle Moksa, Johnson Pang, Diana Mah, Jing Wang, Kiech Fichter, Eric Chuah, Allen Delaney, Rob Kirkpatrick, Agnes Baross, Sarah Barber, Mabel Brown-John, Steve S. Chand, William Chow, Ryan Babakaiff, Dave Wong, Corey Matsuo, Jaclyn Beland, Susan Gibson, Luis delRio, Ruth Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao, Kim MacDonald, Mike R. Mayo, Josh Moran, Diana Palmquist, JR Santos, Duane Smailus, Jeff Stott, Miranda Tsai, George Yang, Jacquie Schein, Asim Siddiqui, Steven Jones, Rob Holt, Marco Marra.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 142 Row: f Column: 24
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 45361358.
Location/Qualifiers
1. .2526
/organism="Xenopus tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="MGC:76140 IMAGE:5335186"
/tissue_type="Embryo, Silurana tropicalis, tailbud (stages 20-27)"
/clone_lib="NICHD_XGC_Emb7"
/lab_host="DH10B"
/notes="Vector: pCMV-SPORT6.1"
1. .2526
/gene="MGC76140"
/db_xref="GeneID:394869"
61. .1056
/gene="MGC76140"
/codon_start=1
/product="hypothetical protein LOC394869"
/protein_id="AAH63914.1"
/db_xref="GI:39645383"
/db_xref="GeneID:394869"
/translation="MANFNKGPAYGLSAEVKNKIAQYDPQVEELRLWIEVTGMII
GENFOGLRDGIILCNILNKLQPSGKKIKNEAKLNHKLGNFKQMDYGMKPHD
IFPANDLFGNKNWTQVTSILVSLAGLAKTRGHTSDVIGKYAEKQKQFGDKWKAG
QSVIGLQMGTKCASQAGMTAYGTRRLHYDPKQTDKPDQDTIISLQMGINKGASQAG
MPAGTRDIYDQKAILQPDVNSTISLQMGTKVASQKMSVGLGRQYDPRYCAAP
TEPIIHNGSQGTGTNGSEISDSYQAEYDPQGEYDPDPRYDHGQSDQGDYD"

Query Match 80.8%; Score 20.2; DB 5; Length 2526;
Best Local Similarity 88.0%; Pred. No. 83;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 AAAAACAACCACTGAGCTGGG 25
||||| ||||| ||||| ||||| |||||
Db 1741 AAAAACAACCACTGAGCTGGG 1717
RESULT 11
BC075515/c
LOCUS
DEFINITION
Xenopus tropicalis hypothetical protein MGC76140, mRNA (cDNA clone MGC:89401 IMAGE:6989687), complete cds.
ACCESSION
BC075515
VERSION
BC075515.1 GI:49522471
KEYWORDS
MGC.
SOURCE
Xenopus tropicalis (Silurana tropicalis)
ORGANISM
Xenopus tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae; Xenopodinae; Xenopus; Silurana.
1 (bases 1 to 2588)
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,J., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettman,M., Madan,A., Rodrigues,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 2588)
Klein,S. and Gerhard,D.S.
Direct Submission
Submitted (28-JUN-2004) National Institutes of Health, Xenopus Gene Collection (XGC), National Institute of Child Health and Human Development, 6100 Executive Boulevard, Room 4B01, Rockville, MD 20892-7510, USA
NIH-MGC Project
Contact: XGC help desk
Email: cgabbs@mail.nih.gov
Tissue Procurement: Robert M. Grainger
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca
Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield, Andy Chan, Steve S. Chand, William Chow, Allison Cloutier, Ruth Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao, Kim MacDonald, Anura Maason, Mike R. Mayo, Josh Moran, Ryan Morin, Teika Olson, Diana Palmquist, Anca Petrescu, Anna Liisa Prahbu, Parvaneh Saeedi, JR Santos, Angelique Schnerch, Ursula Skalska, Duane Smailus, Jeff Stott, Miranda Tsai, George Yang, Jacquie Schein, Asim Siddiqui, Rob Holt, Marco Marra.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 170 Row: g Column: 21


```
misc_feature
gene
CDS
misc_feature
ORIGIN
Query Match 80.8%; Score 20.2; DB 8; Length 114676;
Best Local Similarity 88.0%; Pred. No. 90;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAACCAACCACTGAGCTGGG 25
DB 33952 AAAAAACCAACCACTGAGCTGGG 33976

RESULT 13
AC098690
LOCUS AC098690 Homo sapiens chromosome 1 clone RP11-507I14, complete sequence.
DEFINITION AC098690
ACCESSION AC098690.3 GI:33413356
VERSION AC098690.3
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 128000)
AUTHORS Kaul, R.K., Olson, M.V., Zhou, Y., James, R.A., Rouse, G., Wu, Z.,
Saenphimmachak, C., Buckley, D., Kibukawa, M., Raymond, C. and
Haugen, E.D.
TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 128000)
AUTHORS Kaul, R.K., Olson, M.V., Raymond, C. and Haugen, E.D.
TITLE Direct Submission
JOURNAL Submitted (30-OCT-2001) Genome Center, University of Washington,
Box 352145, Seattle, WA 98195, USA
REFERENCE 3 (bases 1 to 128000)
AUTHORS Kaul, R.K., Olson, M.V., Zhou, Y., James, R.A., Rouse, G., Wu, Z.,
Saenphimmachak, C., Phelps, K.A., Raymond, C. and Haugen, E.D.
TITLE Direct Submission
JOURNAL Submitted (26-APR-2002) Genome Center, University of Washington,
Box 352145, Seattle, WA 98195, USA
REFERENCE 4 (bases 1 to 128000)
AUTHORS Kaul, R.K., Olson, M.V., Zhou, Y., James, R.A., Rouse, G., Wu, Z.,
Saenphimmachak, C., Buckley, D., Kibukawa, M., Raymond, C. and
Haugen, E.D.
TITLE Direct Submission
JOURNAL Submitted (02-AUG-2003) Genome Center, University of Washington,
Box 352145, Seattle, WA 98195, USA
COMMENT On Aug 2, 2003 this sequence version replaced gi:20330880.
-----
Center: University of Washington Genome Center
Center code: UWGC
Web site: http://www.genome.washington.edu
Contact: uwgchgs@u.washington.edu
```

```
/gene="RP11-539G11.1"
/locus_tag="RP11-539G11.1-002"
Join(Complement(AC098691.2:8809..8899),
Complement(102730..102821),Complement(101148..101254),
Complement(101005..101052),Complement(97348..97445),
Complement(94446..94531),Complement(94156..94239),
Complement(93870..94027),Complement(92012..92115),
Complement(91404..91550),Complement(55037..55117),
Complement(54683..54798),Complement(52785..52864),
Complement(51767..51870),Complement(45800..45910),
Complement(44072..44162),Complement(40536..40647),
Complement(40241..40447))
/gene="RP11-539G11.1"
/locus_tag="RP11-539G11.1-002"
/product="novel protein (FLJ36760)"
/note="match: cDNAs: Em:BX649037.1"
Join(Complement(AC098691.2:8809..8839),
Complement(102730..102821),Complement(101148..101254),
Complement(101005..101052),Complement(97348..97445),
Complement(94446..94531),Complement(94156..94239),
Complement(93870..94027),Complement(92012..92115),
Complement(91404..91550),Complement(55037..55117),
Complement(54683..54798),Complement(52785..52864),
Complement(51767..51870),Complement(45800..45910),
Complement(44072..44162),Complement(40536..40647),
Complement(40241..40447))
/gene="RP11-539G11.1"
/locus_tag="RP11-539G11.1-002"
/standard_name="OTTHUMP0000011786"
/codon_start=1
/product="novel protein (FLJ36760)"
/protein_id="CA115306.1"
/db_xref="GI:55962285"
/db_xref="GOA:Q5R113"
/db_xref="InterPro:IPR004179"
/db_xref="UniProt/TREMBL:Q5R113"
/tranlation="MDEGVNFKPTEAGRLMAYIITFETVKKFTYISGKETLSLVLTL
IAGCKEFLDIQRINEKTLNLTNDPNRITIRFPMEGRIKTRKMNKVLCLQAOLGCI
PIQDFALTQDTAKYIPRHSRITRCLTISNAIVNAGLTISFKIETDARELEILNRHP
PFQTQITWYLPYELKVEQITRSDTATLIVLRNFEQLOTRTASDSHYVT
LTIIGADNOVVYLLHTITDVSLLKAGSWAKIAVKRALKSEDLINLISSEFVGLDIOQ
KLTVYLPKPRGNOITWQKSETQIISHKSDISTIAGPNKGTASKKPGNRCNHL
CKSKTCHGDCCKIQAOKSEIKESTISSYSLDLNRNAVSVVPVKRLKIQMNKQSQ
VLDKFGFTPKKSLSESRSSEINSELPIEQWQDQPIYGVKROEPEYQDKEVLNV
NPELGNVWDDFDDLEVTSTDTETKISVLPHMRCRIFLYQILLCPSSVHPP"
EMNSGFVSHSEHSDLSNSAMPFSASMTKLPOQAGNAVIVHFQERKPNLSPEIE
KQCFTFSEKNPNSSNYKKVDFPIRNECKKEVDPR"
Join(Complement(AC098691.2:8809..8839),
Complement(102730..102821),Complement(101148..101254),
Complement(101005..101052),Complement(97348..97445),
Complement(94156..94239),Complement(94446..94531),
Complement(92012..92115),Complement(93870..94027),
Complement(91404..91550),Complement(55037..55117),
Complement(54683..54798),Complement(51767..51870),
Complement(45800..45910),Complement(40571..40647))
/gene="RP11-539G11.1"
/locus_tag="RP11-539G11.1-002"
/standard_name="OTTHUMP0000011785"
/note="match: proteins: Tr:Q8N9Q0"
/codon_start=1
/product="novel protein (FLJ36760)"
/protein_id="CA115307.1"
/db_xref="GI:55962286"
/db_xref="GOA:Q8N9Q0"
/db_xref="InterPro:IPR004179"
/db_xref="UniProt/TREMBL:Q8N9Q0"
/tranlation="MDEGVNFKPTEAGRLMAYIITFETVKKFTYISGKETLSLVLTL
IAGCKEFLDIQRINEKTLNLTNDPNRITIRFPMEGRIKTRKMNKVLCLQAOLGCI
PIQDFALTQDTAKYIPRHSRITRCLTISNAIVNAGLTISFKIETDARELEILNRHP
PFQTQITWYLPYELKVEQITRSDTATLIVLRNFEQLOTRTASDSHYVTIIGADNOVVYLL
HYVSKLEKIGITLSNAIVNAGLTISFKIETDARELEILINRHPFGTQIKETVMYPL
KYELKVEQITRSDTATLIVLRNFEQLOTRTASDSHYVTIIGADNOVVYLLH
KITDSVLLKAGSWAKIAVKRALKSEDLINLISSEFVGLDIOQKLTVYLPKPRGNGN
```

```

----- Project Information
Center project name: chr-1
Center clone name: RP11-507114 (sc0371)
----- Summary Statistics
Sequencing vector: plasmid; L08752; 100% of reads
Chemistry: Dye-terminator ET; 49% of reads
Assembly: Dye-terminator Big Dye; 51% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 127931 bases at least Q40
Consensus quality: 127937 bases at least Q30
Consensus quality: 128000 bases at least Q20
Insert size: 128000; sum-of-contigs
Quality coverage: 19.0x in Q20 bases; sum-of-contigs
-----
Overlapping Sequences:
5': RF4-648113 (UWGC:sc0196) AC093563, 110680-bp overlap
3': RP11-539011, BX323048, 45716-bp overlap
Note: This is a partial submission. The full clone overlaps are
not included.
-----
Sequence Quality Assessment:
This entry has been annotated with sequence quality
estimates computed by the Phrap assembly program.
Quality levels above 40 are expected to have less than
1 error in 10,000 bp. Base-by-base quality values are
not generally visible from the GenBank flat file format
but are available as part of this entry's ASN.1 file.
-----
This sequence was finished as follows unless otherwise noted:
all regions were either double-stranded or sequenced with an
alternate chemistry or covered by high quality data (i.e., Phred
quality >= 30); an attempt was made to resolve all sequencing
problems, such as compressions and repeats; all regions were
covered by at least one plasmid subclone or more than one M13
subclone; and the assembly was confirmed by restriction digest.
-----
Sequence Validation:
This sequence has been validated by Multiple Complete Digest
fingerprinting. Comparison of the experimentally derived digest
fragments with sequence-predicted fragments is given below.
The electronically-digested sequence consists of both insert and
vector, in order to accurately represent the entire circular BAC.
Small fragments below a variable cutoff (approximately 400-800 bp)
are not resolved in the fingerprint and hence do not appear
in the table. There are no significant remaining discrepancies
between the experimental and predicted values. Uniquely ordered
fragments are separated by dashed lines.
BgIII          EcoRI          HindIII
SeqDerMap  FngPrnt  SeqDerMap  FngPrnt  SeqDerMap  FngPrnt
-----
3177      3147      8696      8841      4087      4011
-----
2067      2090      6          6          6382      6489
-----
11746     11739     2995      3061      512       <800
-----
9669      9454      5942      6031      449       <800
-----
1898      1917      3870      3904      1289      1263
-----
450        <800      2692      2709      14        <800
-----
1732      1742      1474      1448      1355      1327
-----
1038      1020      4248      4255      1943      1915
-----
1291      1271      1160      1146      6484      6489
-----
7167      7290      10782     10713     1793      1915
-----

```

4631	4627	1448	1448	1322	1327
1533	1524	5710	5687	7207	7323
1035	1020	3032	3061	12	<800
2039	2090	2692	2709	316	<800
6302	6289	10021	10083	5438	5542
1276	1271	941	942	7483	7323
10795	10817	4920	4957	361	<800
1079	1020	3375	3439	1133	1138
206	<800	9374	9285	1740	1773
4633	4627	2974	2909	181	<800
964	1020	4638	4637	2549	2554
478	<800	6110	6031	7689	7323
680	<800	3061	3061	1548	1556
1433	1424	313	<800	3231	3273
5765	5740	2934	2909	5632	5542
6590	6607	1809	1868	59	<800
9335	9454	5126	4957	354	<800
792	<800	1849	1868	1132	1138
5538	5483	11502	11321	274	<800
3107	3147	9616	9605	1787	1773
973	1020	3438	3439	1643	1616
2139	2090	3061	3061	2023	2112
8116	8094	197	<800	1280	1263
748	<800	13507	13631	338	<800
1910	1917	1893	1868	48	<800
2159	2090	4888	4957	578	<800
2441	2471	354	<800	5807	5542
1938	1917	3986	3904	1891	1915
778	<800			7303	7323
6140	6289			503	<800
2505	2471			100	<800
7054	6985			1419	1416
1435	1424			122	<800
4273	4244			8785	8814
6253	6289			1533	1556
1000	1020			154	<800
2046	2090			48	<800

repeat_region 11213..11247
/rpt_family="Aln"
repeat_region 14788..14814
/rpt_family="(TTA)n"
repeat_region 14815..15576
/rpt_family="L1"
repeat_region 15574..16532
/rpt_family="L1"
repeat_region 16762..17064
/rpt_family="Alu"
repeat_region 20843..22355
/rpt_family="L1"
repeat_region 22359..22696
/rpt_family="L1"
repeat_region 22705..23010
/rpt_family="Alu"
repeat_region 23834..24127
/rpt_family="Alu"
repeat_region 24154..24601
/rpt_family="L1"
repeat_region 25597..26090
/rpt_family="L1"
repeat_region 26229..26249
/rpt_family="AT_rich"
repeat_region 29072..29257
/rpt_family="L1"
repeat_region 29450..29851
/rpt_family="L1"
repeat_region 29841..30049
/rpt_family="L1"
repeat_region 30050..30121
/rpt_family="Alu"
repeat_region 30122..30693
/rpt_family="L1"
repeat_region 30694..30996
/rpt_family="Alu"
repeat_region 30997..31396
/rpt_family="L1"
repeat_region 31397..31700
/rpt_family="Alu"
repeat_region 31701..32174
/rpt_family="L1"
repeat_region 32198..32500
/rpt_family="Alu"
repeat_region 32502..32924
/rpt_family="MaLR"
repeat_region 32939..32994
/rpt_family="A-rich"
repeat_region 34831..35126
/rpt_family="Alu"
repeat_region 35193..35649
/rpt_family="L1"
repeat_region 35863..36526
/rpt_family="ERV1"
repeat_region 36527..36951
/rpt_family="ERV1"
repeat_region 37518..37562
/rpt_family="AT_rich"
repeat_region 37630..37897
/rpt_family="L1"
repeat_region 38642..38967
/rpt_family="ERV1"
repeat_region 39307..39606
/rpt_family="Alu"
repeat_region 40032..40362
/rpt_family="MaLR"
repeat_region 40987..41230
/rpt_family="L1"
repeat_region 41658..41954
/rpt_family="Alu"
repeat_region 42028..42085
/rpt_family="L1"
repeat_region 42086..42361

repeat_region /rpt_family="L1"
42465..43084
/rpt_family="L1"
repeat_region 43084..44825
/rpt_family="L1"
misc_feature 43512..43698
/note="unresolved homopolymeric repeat"

Query Match 80.8%; Score 20.2; DB 8; Length 128638;
Best Local Similarity 88.0%; Pred. No. 90;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAACACACCACTGAGCTGGG 25
|||||
Db 11242 AAAAAACACACCACTCATCTGGG 11266

RESULT 15
AL935194 145569 bp DNA linear VRT 13-DEC-2002
LOCUS Zebrafish DNA sequence from clone CH211-237E12, complete sequence.
DEFINITION
ACCESSION AL935194
VERSION AL935194.4 GI:26985414
KEYWORDS HTG.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 145569)
Giselle,H.
Direct Submission
Submitted (12-DEC-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries: zface@sanger.ac.uk
Clone requests: clonerequest@sanger.ac.uk
On Dec 16, 2002 this sequence version replaced gi:24940082.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: zface@sanger.ac.uk

During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest, except on the rare
occasion of the clone being a YAC.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep Repeat names
beginning 'Dr' were identified by the Recon repeat discovery system
(Zhirong Bao and Sean Eddy, submitted), and those beginning 'drr',
were identified by Rick Waterman (Stephen Johnson lab, WashU). For
further information see http://www/Projects/D_rerio/fishmask.shtml
CH211-237E12 is from a CHORI-211 BAC library
VECTOR: pTARBAC2.1.
Location/Qualifiers
1..145569
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="CH211-237E12"

FEATURES
source

```

ORIGIN
Query Match      80.8%; Score 20.2; DB 5; Length 145569;
Best Local Similarity 88.0%; Pred. No. 90;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 AAAAAACACCACTGAGCTGGG 25
Db 33473 AAAAAACACCACTGAGCTGGG 33497

RESULT 16
AC053517/c
LOCUS AC053517.7 GI:15451722 linear HTG 06-SEP-2001
DEFINITION Homo sapiens chromosome 4 clone RP11-455G12, WORKING DRAFT
FEATURES             Source
ACCESSION AC053517
VERSION AC053517.7 GI:15451722
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
          Hominidae; Homo.
REFERENCE 1 (bases 1 to 155263)
AUTHORS Waterston,R.H.
TITLE The sequence of Homo sapiens clone
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 155263)
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (16-APR-2000) Genome Sequencing Center, Washington
          University School of Medicine, 4444 Forest Park Parkway, St. Louis,
          MO 63108, USA
COMMENT On Sep 6, 2001 this sequence version replaced gi:11192229.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H_NH0455G12
----- Summary Statistics -----
Sequencing vector: M13; 99%
Sequencing vector: plasmid; 0%
Chemistry: Dye-primer ET; 99% of reads
Chemistry: Dye-terminator Big Dye; 0% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 150711 bases at least Q40
Consensus quality: 152938 bases at least Q30
Consensus quality: 154294 bases at least Q20
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 14 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
1 1339: contig of 1339 bp in length
* 1340: gap of unknown length
* 1440: contig of 2429 bp in length
* 3869: gap of unknown length
* 3969: contig of 2989 bp in length
* 6958: gap of unknown length
* 7058: contig of 3693 bp in length
* 10751: gap of unknown length
* 10851: contig of 4158 bp in length
* 10851: gap of unknown length
* 15009: gap of unknown length
* 15109: contig of 5421 bp in length
*
20530: gap of unknown length
* 20630: contig of 4515 bp in length
* 25144: gap of unknown length
* 32639: contig of 7395 bp in length
* 32739: gap of unknown length
* 32740: contig of 8751 bp in length
* 41491: gap of unknown length
* 41591: contig of 14510 bp in length
* 56101: gap of unknown length
* 56201: contig of 15145 bp in length
* 71346: gap of unknown length
* 71446: contig of 18771 bp in length
* 90217: gap of unknown length
* 90317: contig of 19051 bp in length
* 103667: gap of unknown length
* 109468: gap of unknown length
* 155263: contig of 45796 bp in length.
          Location/Qualifiers
            1..155263
              /organism="Homo sapiens"
              /mol_type="genomic DNA"
              /db_xref="taxon:9606"
              /chromosome="4"
              /clones="RP11-455G12"
            1..1339
              /note="assembly_name:Contig68"
            1340..1439
              /estimated_length=unknown
            1440..3868
              /note="assembly_name:Contig69"
              clone_end:SP6
              vector_side:right"
            3869..3968
              /estimated_length=unknown
            3969..6957
              /note="assembly_name:Contig70"
            6958..7057
              /estimated_length=unknown
            7058..10750
              /note="assembly_name:Contig71"
            10751..10850
              /estimated_length=unknown
            10851..15008
              /note="assembly_name:Contig72"
            15009..15108
              /estimated_length=unknown
            15109..20529
              /note="assembly_name:Contig73"
            20530..20629
              /estimated_length=unknown
            20630..25144
              /note="assembly_name:Contig74"
            25145..25244
              /estimated_length=unknown
            25245..32639
              /note="assembly_name:Contig75"
            32640..32739
              /estimated_length=unknown
            32740..41490
              /note="assembly_name:Contig76"
            41491..41590
              /estimated_length=unknown
            41591..56100
              /note="assembly_name:Contig77"
            56101..56200
              /estimated_length=unknown
            56201..71345
              /note="assembly_name:Contig78"
              clone_end:T7
              vector_side:right"
            71346..71445
              /estimated_length=unknown
            71446..90216
              /note="assembly_name:Contig79"

```


misc_feature 4609. .6383 /note="assembly_fragment"
gap 6384. .6483 /estimated_length=100
misc_feature 6484. .8884 /note="assembly_fragment"
gap 8885. .8984 /estimated_length=100
misc_feature 8985. .11538 /note="assembly_fragment"
gap 11539. .11638 /estimated_length=100
misc_feature 11639. .15115 /note="assembly_fragment"
gap 15116. .15215 /estimated_length=100
misc_feature 15216. .18012 /note="assembly_fragment"
gap 18013. .18112 /estimated_length=100
misc_feature 18113. .22355 /note="assembly_fragment"
gap 22356. .22455 /estimated_length=100
misc_feature 22456. .26560 /note="assembly_fragment"
gap 26561. .26660 /estimated_length=100
misc_feature 26661. .30795 /note="assembly_fragment"
gap 30796. .30895 /estimated_length=100
misc_feature 30896. .35255 /note="assembly_fragment"
gap 35256. .35355 /estimated_length=100
misc_feature 35356. .41743 /note="assembly_fragment"
gap 41744. .41843 /estimated_length=100
misc_feature 41844. .82634 /note="assembly_fragment"
gap 82635. .82734 /estimated_length=100
misc_feature 82735. .89758 /note="assembly_fragment"
gap 89759. .89858 /estimated_length=100
misc_feature 89859. .100523 /note="assembly_fragment"
gap 100524. .100623 /estimated_length=100
misc_feature 100624. .111515 /note="assembly_fragment"
gap 111516. .111615 /estimated_length=100
misc_feature 111616. .132549 /note="assembly_fragment"
gap 132550. .132649 /estimated_length=100
misc_feature 132650. .155669 /note="assembly_fragment"
gap 155670. .155769 /estimated_length=100
misc_feature 155770. .159976 /note="assembly_fragment"
clone_end:77
vector_side:right"

ORIGIN

Query Match 80.8%; Score 20.2; DB 14; Length 159976;
Best Local Similarity 88.0%; Pred. No. 91;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAAACACCACTGAGCTGGG 25
|||||
Db 111399 AAAAAAACACCACTGATCTAGG 111375
|||||

RESULT 18

AC073301

LOCUS

DEFINITION

AC073301

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

173256 bp DNA linear HTG 10-MAY-2001

Homo sapiens chromosome 1 clone RP11-221N16 map 1, WORKING DRAFT

SEQUENCE, 23 unordered pieces.

HTG; HTGS_PHASE1; HTGS_DRAFT.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

1 (bases 1 to 173256)

Birren,B., Linton,L., Nusbaum,C. and Lander,E.

Unpublished

2 (bases 1 to 173256)

Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,

Anderson,S., Baldwin,J., Barna,N., Bastien,V., Bada,F.,

Boguslavsky,L., Boukhgalter,B., Brown,A., Burkett,G., Collins,S.,

Compiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,

Collamore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S.,

Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,

Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,

Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,

Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,

Klein,J., Lacroque,K., Lamazares,R., Landers,T., Lehoczy,J.,

Levine,R., Liu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,

McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,

Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,

Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,

O'Neil,D., Olivat,T.M., Oliver,J., Peterson,K., Pierre,N.,

Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,

Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,

Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,

Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,

Vassiliev,H., Viel,R., Vo.A., Wilson,B., Wu.X., Wyman,D., Ye,W.J.,

Young,G., Zainoun,J., Zimmer,A. and Zody,M.

Direct Submission

Submitted (14-JUN-2000) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

On May 10, 2001 this sequence version replaced gi:8516089.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L7730

Center clone name: 221_N_16

----- Summary Statistics

Sequencing vector: M13: M77815; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 160406 bases at least Q40

Consensus quality: 166935 bases at least Q30

Consensus quality: 169615 bases at least Q20

Insert size: 163000; agarose-fp

Insert size: 171056; sum-of-contigs

Quality coverage: 4.3 in Q20 bases; agarose-fp

Quality coverage: 4.1 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently


```

AC013557/c
LOCUS       AC013557               178806 bp      DNA      linear      HTG 24-AUG-2002
DEFINITION  Homo sapiens chromosome 11 clone RP11-28D2 map 11, WORKING DRAFT
AC013557
SEQUENCE    4 unordered pieces.
AC013557
ACCESSION   AC013557.3  GI:7107807
VERSION     HTG; HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS    Homo sapiens (human)
SOURCE      Homo sapiens
ORGANISM    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominoidea; Homo.
REFERENCE   1 (bases 1 to 178806)
AUTHORS     Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE       Homo sapiens chromosome 11, clone RP11-28D2
JOURNAL     Unpublished
REFERENCE   2 (bases 1 to 178806)
AUTHORS     Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
            Baldwin,J., Barna,N., Beckerly,R., Boguslavkiy,L., Boukhgalter,B.,
            Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A.,
            Cooke,P., DeArelano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,
            Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D.,
            Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
            Lehoczy,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,
            McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrim,J.,
            Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
            Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
            Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
            Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
            Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
DIRECT SUBMISSION
Submitted (13-NOV-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 178806)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,
Barna,N., Bastien,V., Boguslavkiy,L., Boukhgalter,B., Brown,A.,
Camarata,J., Campopiano,A., Chang,J., Choepel,Y., Colangelo,M.,
Collins,S., Collymore,A., Cooke,P., DeArelano,K., Dewar,K.,
Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Hagos,B., Heaford,A., Horton,L., Hulme,W.,
Iliev,I., Johnson,R., Jones,C., Karatas,A., LaRoque,K.,
Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G.,
Maclean,C., Macdonald,P., Marquis,N., Matthews,C., McCarthy,M.,
McEwan,P., McKernan,K., McPheeters,R., Meldrim,J., Meneus,L.,
Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Norbu,C.,
Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J.,
Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C.,
Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,
Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S.,
Severy,P., Sougnuez,C., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S.,
Theodore,J., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
DIRECT SUBMISSION
Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Feb 28, 2000 this sequence version replaced gi:6479117.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L2252
Center clone name: 28 D 2
----- Summary Statistics
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 170392 bases at least Q40
Consensus quality: 176132 bases at least Q30
Consensus quality: 177464 bases at least Q20
Insert size: 176000; agarose-fp
Insert size: 178506; sum-of-contigs
Quality coverage: 6.7 in Q20 bases; agarose-fp
Quality coverage: 6.6 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 1239: contig of 1239 bp in length
* 1240 1339: gap of 100 bp
* 1340 23256: contig of 21917 bp in length
* 23257 23356: gap of 100 bp
* 23357 88091: contig of 64735 bp in length
* 88092 88191: gap of 100 bp
* 88192 178806: contig of 90615 bp in length.
*
* Location/Qualifiers
* 1..178806
* /organism="Homo sapiens"
* /mol_type="genomic DNA"
* /db_xref="taxon:9606"
* /chromosome="11"
* /map="11"
* /clone="RP11-28D2"
* /clone_lib="RPC1-11 Human Male BAC"
* 1..1239
* /note="assembly_fragment"
* 1240..1339
* /estimated_length=100
* 1340..23256
* /note="assembly_fragment"
* clone_end:T7
* vector_side:right"
* 23257..23356
* /estimated_length=100
* 23357..88091
* /note="assembly_fragment"
* clone_end:SP6
* vector_side:left"
* 88092..88191
* /estimated_length=100
* 88192..178806
* /note="assembly_fragment"
*
* ORIGIN
* Query Match 80.8%; Score 20.2; DB 14; Length 178806;
* Best Local Similarity 88.0%; Pred No. 91;
* Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
*
* QY 1 AAAAAAACACCACTGAGCTGGG 25
* | | | | | | | | | | | | | | | | | | | |
* Db 8658 AAAAAAACACCACTGAGCTGGG 8634
*
* RESULT 20
* AC104828
* LOCUS AC104828 181489 bp DNA linear PRI 29-MAY-2002
* DEFINITION Homo sapiens BAC clone RP11-804N11 from 4, complete sequence.
* ACCESSION AC104828
* VERSION AC104828.4 GI:20429638
* KEYWORDS HTG.
* SOURCE Homo sapiens (human)
* ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

REFERENCE

AUTHORS

1 (bases 1 to 181489)
Sulston, J.E. and Waterston, R.

TITLE

Toward a complete human genome sequence

JOURNAL

PUBMED

9847074
Genome Res. 8 (11), 1097-1108 (1998)

REFERENCE

AUTHORS

2 (bases 1 to 181489)

TITLE

Scott, K., Haglund, K. and Doebber, A.

JOURNAL

PUBMED

Unpublished (2001)

REFERENCE

AUTHORS

JOURNAL

PUBMED

3 (bases 1 to 181489)
Waterston, R.H.

TITLE

JOURNAL

PUBMED

Direct Submission
Submitted (21-DEC-2001) Genome Sequencing Center, Washington

REFERENCE

AUTHORS

JOURNAL

PUBMED

University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA

REFERENCE

AUTHORS

JOURNAL

PUBMED

4 (bases 1 to 181489)
Waterston, R.H.

TITLE

JOURNAL

PUBMED

Direct Submission
Submitted (16-FEB-2002) Genome Sequencing Center, Washington

REFERENCE

AUTHORS

JOURNAL

PUBMED

University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA

REFERENCE

AUTHORS

JOURNAL

PUBMED

5 (bases 1 to 181489)
Waterston, R.H.

TITLE

JOURNAL

PUBMED

Direct Submission
Submitted (03-MAY-2002) Genome Sequencing Center, Washington

REFERENCE

AUTHORS

JOURNAL

PUBMED

University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA

REFERENCE

AUTHORS

JOURNAL

PUBMED

6 (bases 1 to 181489)
Waterston, R.H.

TITLE

JOURNAL

PUBMED

Direct Submission
Submitted (29-MAY-2002) Department of Genetics, Washington

REFERENCE

AUTHORS

JOURNAL

PUBMED

University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On May 3, 2002 this sequence version replaced gi:18693591.

COMMENT

JOURNAL

PUBMED

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-703H20; the clone sequenced to the right is RP11-556N4. Actual start of this clone is at base position 1995 of RP11-703H20; actual end is at base position 181489 of RP11-804N11.

The region from 97601 to 97728 is covered only by a per product from clone DNA. Data from AC053517 was used to finish AC104828.

FEATURES

SOURCE

Location/Qualifiers

1. .181489

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="4"

/map="4"

/clone="RP11-804N11"

/clone_lib="RPCI-11"

7. .107

/rpt_family="L1"

107. .405

/rpt_family="L1"

406. .713

/rpt_family="Alu"

714. .906

/rpt_family="L1"

946. .1295

/rpt_family="Alu"

1543. .1844

/rpt_family="Alu"

2424. .2718

/rpt_family="Alu"

4417. .4456

/rpt_family="AT_rich"

4975. .5000

/rpt_family="(CA)n"

5184. .5213

/rpt_family="AT_rich"

5276. .5572

/rpt_family="Alu"

5774. .6138

/rpt_family="L2"

6271. .6296

/rpt_family="AT_rich"

6303. .6624

/rpt_family="Alu"

6759. .7055

/rpt_family="Alu"

7036. .7156

/rpt_family="(GGAA)n"

8574. .8600

/rpt_family="(CA)n"

8977. .9022

/rpt_family="L2"

9131. .9160

/rpt_family="AT_rich"

9327. .9350

/rpt_family="AT_rich"

9463. .10389

/rpt_family="MERI_type"

10390. .10684

/rpt_family="Alu"

10685. .10860

/rpt_family="MERI_type"

11528. .11556

/rpt_family="AT_rich"

11575. .11868

/rpt_family="Alu"

11918. .12357

/rpt_family="L1"

12434. .12458

/rpt_family="(TTTTA)n"

12566. .12755

/rpt_family="L1"

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Woon, P.Y., Zhao, B., Frengen, E., Tateno, M., Catanese, J.J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. *Genomics* 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>

VECTOR: pBACE3.6

```

repeat_region 13279. .13587
/rpt family="Alu"
repeat_region 13630. .13931
/rpt family="Alu"
repeat_region 14094. .14409
/rpt family="Alu"
repeat_region 14559. .14838
/rpt family="Alu"
repeat_region 15975. .16123
/rpt family="Alu"
repeat_region 16182. .16466
/rpt family="Alu"
repeat_region 16467. .16491
/rpt family="CA)n"
repeat_region 16768. .16989
/rpt family="MER2_type"
repeat_region 17327. .17622
/rpt family="Alu"
repeat_region 18500. .18604
/rpt family="L2"
repeat_region 18932. .19222
/rpt family="Alu"
repeat_region 19233. .19534
/rpt family="Alu"
repeat_region 19822. .20133
/rpt family="Alu"
repeat_region 20435. .20716
/rpt family="Alu"
repeat_region 20717. .20786
/rpt family="TA)n"
repeat_region 21032. .21341
/rpt family="Alu"
repeat_region 21342. .21498
/rpt family="MIR"
repeat_region 21794. .21823
/rpt family="AT_rich"
repeat_region 21918. .21962
/rpt family="TATG)n"
repeat_region 22071. .22191
/rpt family="MIR"
repeat_region 22197. .22286
/rpt family="L2"
repeat_region 22287. .22590
/rpt family="Alu"
repeat_region 22591. .22686
/rpt family="L2"

Query Match      80.8%; Score 20.2; DB 8; Length 181489;
Best Local Similarity 88.0%; Pred. No. 91;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AAAAAACACCACTGAGCTGGG 25
    ||||| ||||| ||||| ||||| |||||
Db 86906 AAAAAACACCACTGATCTAGG 86930

RESULT 21
AC099068 Rattus norvegicus clone CH230-7F9, WORKING DRAFT SEQUENCE, 5 linear HTG 09-MAY-2003
DEFINITION
ACCESSION AC099068
VERSION AC099068.9 GI:30467914
KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT; HTGS FULLTOP.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE 1 (bases 1 to 243279)
Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,

```

Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F., Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M., Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E., Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A., Chacko,J., Chavez,D., Chen,G., Chen,Y., Chen,Z., Chu,J., Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L., Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D., Delgado,O., Denson,S., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K., Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G., Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P., Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M., Gregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W., Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K., Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J., Hernandez,R., Hines,S., Hladun,S.B., Hodgson,A., Hognes,M., Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A., Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A., Karpathy,S., Kelly,S., Khan,Z., King,L., Kovar,C., Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J., Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J., Lorensunewa,L., Louisedge,H., Lozado,R.J., Lu,X., Ma,J., Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A., Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E., Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E., Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S., Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L., Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Nair,L., Nwackeleme,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K., Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C., Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L., Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F., Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J., Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H., Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D., Snead,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J., Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Umani,K., Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,R., Warren,R., Wei,X., White,F., Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S., Yen,L., Yoon,L., Yoon,V., Yu,P., Zhang,J., Zhou,X., Zhou,S., Zhao,S., Dunn,D., von Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O., Weinstock,G. and Gibbs,R.A.

Direct Submission
Unpublished
2 (bases 1 to 243279)
Worley,K.C.
Direct Submission
Submitted (09-NOV-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (09-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On May 9, 2003 this sequence version replaced gi:24941198.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

ORGANISM

```

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GCEH
Center clone name: CH230-7F9
----- Summary Statistics
Assembly program: Atlas
Consensus quality: 237282 bases at least Q40
Consensus quality: 238711 bases at least Q30
Consensus quality: 239637 bases at least Q20
Estimated insert size: 245767; sum-of-contigs estimation
Quality coverage: 11x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 5 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved
*
* 1 231160: contig of 231160 bp in length
* 231161 231260: gap of unknown length
* 231261 236392: contig of 5132 bp in length
* 236393 236492: gap of unknown length
* 236493 240099: contig of 3607 bp in length
* 240100 240199: gap of unknown length
* 240200 241548: gap of 1349 bp in length
* 241549 241648: gap of unknown length
* 241649 243279: contig of 1631 bp in length.

FEATURES             Location/Qualifiers
     source            1..243279
                     /organism="Rattus norvegicus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10116"
                     /clone="CH230-7F9"
     misc_feature      1..1310
                     /notes="wgs contig"
     gap              231161..231260
                     /estimated_length=unknown
     misc_feature      231261..232384
                     /notes="wgs contig"
     gap              236393..236492
                     /estimated_length=unknown
     gap              240100..240199
                     /estimated_length=unknown
     gap              241549..241648
                     /estimated_length=unknown

ORIGIN
Query Match      80.88; Score 20.2; DB 14; Length 243279;
Best Local Similarity 88.08; Pred. No. 91;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAACACACACCACTGAGCTGGG 25
Db 137647 AAAAAAAAAACCACTGTGAGCTGGG 137671

RESULT 22
AC095147/c
LOCUS            AC095147            256914 bp    DNA    linear    HTG 09-MAY-2003
DEFINITION      Rattus norvegicus clone CH230-8H22, WORKING DRAFT SEQUENCE, 3
unordered pieces.
ACCESSION      AC095147
VERSION        AC095147.6 GI:30467274
KEYWORDS       HTG; HTGS PHASE1; HTGS DRAFT; HTGS_FULLTOP.
SOURCE         Rattus norvegicus (Norway rat)

```

ORGANISM

Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Rattus.

1 (bases 1 to 256914)
Muzny, D. Marie., Metzker, M. Lee., Abranzon, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Sacoto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensuhsuwa, L., Louised, H., Lozado, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhiney, S., McLeod, M. P., McNeill, T. Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwakoelameh, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Poinexter, A., Popovic, D., Primus, E., Pu, L., L., Puaro, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smajs, D., Speed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Soza, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wlaczky, R., Woodson, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstock, G. and Gibbs, R. A.

Direct Submission
Unpublished
2 (bases 1 to 256914)
Worley, K. C.
Direct Submission
Submitted (16-SEP-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 256914)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (09-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On May 9, 2003 this sequence version replaced gi:24817816.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled with Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas

assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated and oriented. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GTC

Center clone name: CH230-8H22

----- Summary Statistics

Assembly program: Atlas;

Consensus quality: 205127 bases at least Q40

Consensus quality: 208686 bases at least Q30

Consensus quality: 210795 bases at least Q20

Estimated insert size: 213188; sum-of-contigs estimation

Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 253811: contig of 253811 bp in length
* 253812 253911: gap of unknown length
* 253912 255439: contig of 1528 bp in length
* 255440 255539: gap of unknown length
* 255540 256914: contig of 1375 bp in length.

Location/Qualifiers

1. .256914
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-8H22"

258. .925

misc_feature

/note="clone boundary
clone end:Sp6
site:ECORI

end_sequence:BH345121"

complement(245885..252580)

/note="clone boundary

clone end:T7

site:ECORI

end_sequence:BH345119"

complement(251898..252621)

/note="clone boundary

clone end:T7

site:ECORI

end_sequence:BH345119"

253812..253911

/estimated_length=unknown

255440..255539

/estimated_length=unknown

ORIGIN

Query Match 80.8%; Score 20.2; DB 14; Length 256914;

Best Local Similarity 88.0%; Pred. No. 91;

Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAACACCACTGAGCTGGG 25

|||||

Db 60736 AAAAAACACCACTGAGCAAGG 60712

RESULT 23

AC160578

LOCUS

DEFINITION

AC160578

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Bos taurus

Bos taurus

Bos taurus

REFERENCE

AUTHORS

AC160578 290052 bp DNA linear HTG 01-JUL-2005
Bos taurus clone CH240-84A20, *** SEQUENCING IN PROGRESS ***, 48
unordered pieces.

AC160578 2 GI:58300415

HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.

Bos taurus (cow)

Bos taurus

Bos taurus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
Pecora; Bovidae; Bovinae; Bos.

1 (bases 1 to 290052)

Muzny,D.M., Metzker,M.Lee., Abramson,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Alabrooks,S., Amin,A., Anguiano,D.,
Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,P.,
Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
Cleaveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
Gebregeorgis,E., Gear,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
Harvey,Y., Haviak,P., Hawes,A., Henderson,N., Hernandez,J.,
Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
Karpachy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
Lorensuhewa,L., Loulsegied,H., Lozado,R.J., Lu,X., Ma,J.,
Maheshwari,M., Mahindaratne,M., Mahmoud,M., Malloy,K., Mangum,A.,
Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
Mawhney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
Morgan,M., Morris,K., Morris,S., Mundaya,M., Murphy,S., Nair,L.,
Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
Nwackelmehe,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
Pasternak,S., Paul,H., Perez,A., Perez,L., Pfankoch,C.,
Plopper,P., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
Puzo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.,
Sanders,W., Savery,G., Scherer,S., Scott,G., Shatman,S., Shen,H.,
Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajls,D.,
Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Umanai,K.,
Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
Weinstock,G. and Gibbs,R.A.

Direct Submission

Unpublished

2 (bases 1 to 290052)

Worley,K.C.

Direct Submission

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

Submitted (28-APR-2005) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

REFERENCE
AUTHORS
TITLE
JOURNAL

3 (bases 1 to 290052)
Dow Genome Sequencing Consortium.
Direct Submission
Submitted (01-JUL-2005) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

COMMENT

The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: FFGH
Center clone name: CH240-84A20
----- Summary Statistics

Assembly program: Atlas 3.0;
Consensus quality: 268871 bases at least Q40
Consensus quality: 273631 bases at least Q30
Consensus quality: 277456 bases at least Q20
Estimated insert size: 274073; sum-of-contigs estimation
Quality coverage: 5x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)
* NOTE: This sequence may represent more than one clone.
* NOTE: This is a 'working draft' sequence. It currently
consists of 48 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.

1
7000: contig of 7000 bp in length
7001: 7050: gap of 50 bp
7051: 9676: contig of 2626 bp in length
9676: 9776: gap of unknown length
9776: 16485: contig of 6709 bp in length
16485: 16535: gap of 50 bp
16535: 17680: contig of 1145 bp in length
17680: 22053: gap of 4373 bp
22053: 24414: contig of 2361 bp in length
24414: 24464: gap of 50 bp
24464: 30888: contig of 6424 bp in length
30888: 30938: gap of 50 bp
30938: 40084: contig of 9146 bp in length
40084: 40134: gap of 50 bp
40134: 45676: contig of 5542 bp in length
45676: 45726: gap of 50 bp
45726: 57298: contig of 11572 bp in length
57298: 57299: gap of 50 bp
57299: 64857: contig of 7509 bp in length
64857: 64907: gap of 50 bp
64907: 70695: contig of 5788 bp in length
70695: 70795: gap of unknown length
70795: 77607: contig of 6812 bp in length
77607: 77657: gap of 50 bp
77657: 81575: contig of 3918 bp in length
81575: 81675: gap of unknown length
81675: 85723: contig of 4048 bp in length
85723: 81676

* 85724 85773: gap of 50 bp
85773: 93143: contig of 7370 bp in length
93143: 93193: gap of 50 bp
93193: 94931: contig of 1738 bp in length
94931: 94981: gap of 50 bp
94981: 112800: contig of 17819 bp in length
112800: 112801: gap of 50 bp
112801: 112851: contig of 11325 bp in length
112851: 115976: contig of 50 bp
115976: 116025: gap of 50 bp
116025: 121069: contig of 5044 bp in length
121069: 121119: gap of 50 bp
121119: 126466: contig of 5347 bp in length
126466: 126516: gap of 50 bp
126516: 142203: contig of 15687 bp in length
142203: 142303: gap of unknown length
142303: 147764: contig of 5461 bp in length
147764: 147814: gap of 50 bp
147814: 152013: contig of 4199 bp in length
152013: 152063: gap of 50 bp
152063: 157204: contig of 5141 bp in length
157204: 157254: gap of 50 bp
157254: 209028: contig of 51774 bp in length
209028: 209078: gap of 50 bp
209078: 212455: contig of 3377 bp in length
212455: 212995: gap of 540 bp
212995: 214790: contig of 1795 bp in length
214790: 214840: gap of 50 bp
214840: 216146: contig of 1306 bp in length
216146: 216246: gap of unknown length
216246: 226185: contig of 9939 bp in length
226185: 226235: gap of 50 bp
226235: 233347: contig of 7112 bp in length
233347: 233447: gap of unknown length
233447: 235721: contig of 2274 bp in length
235721: 235771: gap of 50 bp
235771: 237540: contig of 1769 bp in length
237540: 237640: gap of unknown length
237640: 246873: contig of 9233 bp in length
246873: 246923: gap of 50 bp
246923: 252361: contig of 5438 bp in length
252361: 252411: gap of 50 bp
252411: 255888: contig of 3477 bp in length
255888: 255938: gap of 50 bp
255938: 260942: contig of 5004 bp in length
260942: 261042: gap of unknown length
261042: 262102: contig of 1060 bp in length
262102: 262202: gap of unknown length
262202: 263692: contig of 1490 bp in length
263692: 263792: gap of unknown length
263792: 264924: contig of 1032 bp in length
264924: 264925: gap of unknown length
264925: 266484: contig of 1560 bp in length
266484: 266584: gap of unknown length
266584: 267818: contig of 1234 bp in length
267818: 267918: gap of unknown length
267918: 269286: contig of 1368 bp in length
269286: 269386: gap of unknown length
269386: 270427: contig of 1041 bp in length
270427: 269387

Query Match 80.8%; Score 20.2; DB 14; Length 290052;
Best Local Similarity 88.0%; Pred. No. 92;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAACACCACTGAGCTGGG 25

DB 199963 AAAACACACCACTGAGCTGTG 199987

RESULT 24
CR854901/c
LOCUS

CR854901 192168 bp DNA linear HTG 19-MAY-2005
Danio rerio clone DK5Y-5M2, *** SEQUENCING IN PROGRESS ***, 11
unordered pieces.

```
ACCESSION CR854901
VERSION GI:66351706
KEYWORDS HTG; HTGS PHASE1.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
          Cypriniformes; Cyprinidae; Danio.
          1 (bases 1 to 192168)
REFERENCE Sims,S.
AUTHORS Direct Submission
TITLE Submitted (17-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
JOURNAL Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
          zfsh-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
          On May 20, 2005 this sequence version replaced gi:54695370.
COMMENT ----- Genome Center
          Center: Wellcome Trust Sanger Institute
          Center code: SC
          Web site: http://www.sanger.ac.uk
          Contact: zfsh-help@sanger.ac.uk
          ----- Project Information
          Center project name: zK5M2
          ----- Summary Statistics
          Chemistry: Dye-terminator; 100% of reads
          Consensus quality: 190046 bases at least Q40
          Consensus quality: 190568 bases at least Q30
          Consensus quality: 190823 bases at least Q20
          Insert size: 191168; sum-of-contigs
          Insert size: 191878; 5.8% error; agarose-fp
          Quality coverage: 6.16x in Q20 bases; sum-of-contigs Quality
          coverage: 6.35x in Q20 bases; agarose-fp
          -----
          * NOTE: This is a 'working draft' sequence. It currently
          * consists of 11 contigs. The true order of the pieces
          * is not known and their order in this sequence record is
          * arbitrary. Gaps between the contigs are represented as
          * runs of N, but the exact sizes of the gaps are unknown.
          * This record will be updated with the finished sequence.
          * as soon as it is available and the accession number will
          * be preserved.
          *
          * 1 2172: contig of 2172 bp in length
          * 2173 2272: gap of 100 bp
          * 46258: contig of 4386 bp in length
          * 46359 52193: contig of 5835 bp in length
          * 52194 52293: gap of 100 bp
          * 52294 55247: contig of 2954 bp in length
          * 55248 55347: gap of 100 bp
          * 55348 62364: contig of 7017 bp in length
          * 62365 62465: gap of 100 bp
          * 62465 68493: contig of 6028 bp in length
          * 68493 88001: contig of 19409 bp in length
          * 88002 88101: gap of 100 bp
          * 88102 107897: contig of 19796 bp in length
          * 107898 135043: contig of 27046 bp in length
          * 135044 152062: contig of 100 bp
          * 152063 152162: contig of 16919 bp in length
          * 152063 192168: gap of 100 bp
          * 152163 192168: contig of 40006 bp in length.
          *
          * Location/Qualifiers
          1..192168
             /organism="Danio rerio"
             /mol_type="genomic DNA"
             /db_xref="taxon:7955"
             /clone="PKEY.5M2"
             /clone_lib="DanioKey"
             1..2172
             /note="assembly fragment:00037
             fragment_chain:1"
             2273..46258
             misc_feature
             misc_feature

/note="assembly_fragment:01455
fragment_chain:1"
46359..52193
/note="assembly_fragment:00049
fragment_chain:1"
52294..55247
/note="assembly_fragment:00027
fragment_chain:2"
55348..62364
/note="assembly_fragment:00161
fragment_chain:2"
62465..68492
/note="assembly_fragment:00099
fragment_chain:2"
68593..88001
/note="assembly_fragment:00366
fragment_chain:2"
88102..107897
/note="assembly_fragment:00555
fragment_chain:2"
10798..135043
/note="assembly_fragment:00752
fragment_chain:2"
135144..152062
/note="assembly_fragment:00233
fragment_chain:2"
152163..192168
/note="assembly_fragment:01010
fragment_chain:2"
clone_end:SP6
vector_side:right"

ORIGIN
Query Match 79.2%; Score 19.8; DB 14; Length 192168;
Best Local Similarity 91.3%; Pred.No. 1.4e+02;
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAACAACCACTGAGCTG 23
   |||||
Db 80772 AACAAACCACTGAGCTG 80750

RESULT 25
BX469934
LOCUS Zebrafish DNA sequence from clone CH211-209P5, complete sequence.
DEFINITION BX469934
ACCESSION BX469934
VERSION BX469934.5 GI:35760789
KEYWORDS HTG.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
          Cypriniformes; Cyprinidae; Danio.
          1 (bases 1 to 225586)
REFERENCE Wood,J.
AUTHORS Direct Submission
TITLE Submitted (25-SEP-2003) Wellcome Trust Sanger Institute, Hinxton,
JOURNAL Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
          zfsh-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
          On Sep 25, 2003 this sequence version replaced gi:35209448.
COMMENT ----- Genome Center
          Center: Wellcome Trust Sanger Institute
          Center code: SC
          Web site: http://www.sanger.ac.uk
          Contact: zfsh-help@sanger.ac.uk
          -----
          During sequence assembly data is compared from overlapping clones.
          Where differences are found these are annotated as variations
          together with a note of the overlapping clone name. Note that the
          variation annotation may not be found in the sequence submission
          corresponding to the overlapping clone, as we submit sequences with
          only a small overlap as described above.
```

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at

http://www.sanger.ac.uk/projects/C_elegans/wormpep Clone-derived Zebrafish pUC subclones occasionally display inconsistency over the length of mononucleotide A/T runs and conserved TA repeats. Where this is found the longest good quality representation will be submitted.

Repeat names beginning 'Dr' were identified by the Recon repeat discovery system (Zhiyong Bao and Sean Eddy, submitted), and those beginning 'drr' were identified by Rick Waterman (Stephen Johnson lab, WashU). For further information see

http://www.sanger.ac.uk/projects/D_rerio/fishmaek.shtml CH211-209P5 is from a CHORI-211 BAC library

VECTOR: pTARBAC2.1.

FEATURES

source

1..225586
Location/Qualifiers
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="CH211-209P5"
/clone_lib="CHORI-211"

ORIGIN

Query Match 79.2%; Score 19.8; DB 5; Length 225586;

Best Local Similarity 91.3%; Pred. No. 1.4e+02;

Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AAAAACAACCACTGAGCTG 23

Db 161991 ACAAACAACCACTGAGCTG 162013

RESULT 26

AC098538 255707 bp DNA linear HTG 10-MAY-2003

LOCUS Rattus norvegicus clone CH230-92N23, *** SEQUENCING IN PROGRESS

DEFINITION ***, 11 unordered pieces.

AC098538 AC098538.6 GI:30521657

KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT; HTGS_ENRICHED.

SOURCE Rattus norvegicus (Norway rat)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

Sciurognathi; Muridae; Murinae; Rattus.

1 (bases 1 to 255707)

Muzny,D,Marie., Metzker,M, Lee., Abramson,S., Adams,C., Alder,J.,

Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,

Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,

Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,

Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,

Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,

Cardenas,V., Carter,K., Cavazos,I., Cesar,H., Center,A.,

Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,

Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cres,A., D'Souza,L.,

Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,

Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,

Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,

Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,

Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,

Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,

Gabregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,

Gunaratanne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,

Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
Hernandez,B., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
Hollins R., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
Kowals,C., Kratt,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
Lorensuhewa,L., Loulseghe,H., Lozado,R.J., Lu,X., Ma,J.,
Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
Mawhinney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
Miloavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
Morgan,M., Morris,K., Morris,S., Muidasa,M., Murphy,M., Nair,L.,
Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
Nwaokemele,O., Okwuonu,G., Olarnpusagoon,A., Pal,S., Parks,K.,
Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkuch,C.,
Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.,
Puaio,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajic,D.,
Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
Steinle,M., Strong,R., Sutton,A., Svatek,A., Taber,P., Taylor,C.,
Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
Valas,R., Vera,V., Villalana,D., Waldron,L., Walker,B., Wang,J.,
Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
Williams,G., Willson,R., Wlarczyk,R., Wooden,H., Worley,K.,
Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
Weinstock,G. and Gibbs,R.A.

Direct Submission

Unpublished

2 (bases 1 to 255707)

Worley,K.C.

Direct Submission

Submitted (24-OCT-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 255707)

Rat Genome Sequencing Consortium.

Direct Submission

Submitted (10-MAY-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

On May 10, 2003 this sequence version replaced gi:25093373.

The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GHX

Center clone name: CH230-92N23

----- Summary Statistics

Assembly program: Atlas 3.0;

Consensus quality: 213914 bases at least Q40

Consensus quality: 218379 bases at least Q30

Consensus quality: 221566 bases at least Q20

Estimated insert size: 232388; sum-of-contigs estimation

Quality coverage: 5x in Q20 bases; sum-of-contigs estimation

*** NOTE: Estimated insert size may differ from sequence length (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently consists of 11 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

* 1 79107: contig of 79107 bp in length
* 79108 79207: gap of unknown length
* 79208 85284: contig of 6077 bp in length
* 85284 85384: gap of unknown length
* 85384 187331: contig of 101947 bp in length
* 187331 187431: gap of unknown length
* 187431 234684: contig of 47252 bp in length
* 234684 234784: gap of unknown length
* 234784 235975: contig of 1192 bp in length
* 235975 236075: gap of unknown length
* 236075 237585: contig of 1509 bp in length
* 237585 237684: gap of unknown length
* 237684 239262: contig of 1578 bp in length
* 239262 239362: gap of unknown length
* 239362 241904: contig of 2542 bp in length
* 241904 242004: gap of unknown length
* 242004 247701: contig of 5697 bp in length
* 247701 247801: gap of unknown length
* 247801 251585: contig of 3784 bp in length
* 251585 251686: gap of unknown length
* 251686 255707: contig of 4022 bp in length.

FEATURES

source
1..255707
Location/Qualifiers
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-92N23"
79108..79207
/estimated_length=unknown
81675..85284
/note="wgs contig"
85285..85384
/estimated_length=unknown
187332..187431
/estimated_length=unknown
187432..188989
/note="wgs contig"
234684..234783
/estimated_length=unknown
235976..236075
/estimated_length=unknown
237585..237684
/estimated_length=unknown
239263..239362
/estimated_length=unknown
241905..242004
/estimated_length=unknown
247702..247801
/estimated_length=unknown
251586..251685
/estimated_length=unknown

ORIGIN

Query Match 79.2%; Score 19.8; DB 14; Length 255707;
Best Local Similarity 91.3%; Pred. No. 1.4e+02;
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 2 AAAAAACACCACTGAGCTGG 24
|||||
Db 223861 AAAAAACATCACCACTGGCTGG 223883
|||||

RESULT 27

CHKHPXPX
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
CHKHPXPX 1288 bp DNA linear VRT 19-JUL-1995
Gallus gallus hypothetical exon, complete cds.
M84354
M84354.1 GI:211948
homeobox protein.
Gallus gallus (chicken)
Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.

REFERENCE
AUTHORS
TITLE
JOURNAL
PUBMED
COMMENT
FEATURES
1 (bases 1 to 1288)
Goldberg, G.S. and Kaczmarczyk, W.
Sequence of a novel chicken genomic DNA fragment that hybridizes to the murine Hox-3.1 homeobox
Gene 121 (2), 397-398 (1992)
1359990
Original source text: Gallus gallus DNA.
Location/Qualifiers
1..1288
/organism="Gallus gallus"
/mol_type="genomic DNA"
/db_xref="taxon:9031"
21..25
/note="putative"
306..316
/note="putative"
/rpt_type=tandem
343..353
/note="putative"
/rpt_type=tandem
372..380
/note="putative"
384..758
/note="hypothetical exon homologous to murine Hox-3.1 homeobox; putative"
/codon_start=1
/protein_id="AAA70193.1"
/db_xref="GI:211949"
/translation="MSACALPARVGLCKWASGPSRTAAAEYRPNRYVSLPABLDPATY
DTPLKRRABERLAIRALKRQYLQLNTPKPRVIVSGAGGAGGLWASPPHTDGS
GPRQSLSPSRGLAGGAVFLSL"
384..754
/note="putative"
755..799
/note="putative"
800..965
/note="putative"
1279..1284
/note="putative"

TATA_signal

repeat_unit

repeat_unit

GC_signal

CDS

exon

intron

exon

polyA_signal

ORIGIN

Query Match 77.6%; Score 19.4; DB 5; Length 1288;
Best Local Similarity 95.2%; Pred. No. 2.1e+02;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 AAAAAACACCACTGAGC 21
|||||
Db 196 AAAAAACACCACTCTGAGC 216
|||||

RESULT 28

GGWALTER
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
GGWALTER 1288 bp DNA linear VRT 17-FEB-1997
G.gallus DNA for hypothetical protein walter.
X60778
X60778.1 GI:517249
Gallus gallus (chicken)
Gallus gallus

RESULT 32
AC153907/c

AL953906 172246 bp DNA linear ROD 31-JAN-2003
Mouse DNA sequence from clone RP23-109E21 on chromosome 2, complete
sequence.

ACCESSION AL953906
VERSION AL953906.6 GI:28193313
KEYWORDS HTG.

Mus musculus (house mouse)
Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 172246)

Chapman, J.

Direct Submission
Submitted (31-JAN-2003) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk

On Feb 1, 2003 this sequence version replaced GI:27902085.
Sequence from the Mouse Genome Sequencing Consortium whole genome
shotgun may have been used to confirm this sequence. Sequence data
from the whole genome shotgun alone has only been used where it has
a phred quality of at least 30.

----- Genome Center

Center: Wellcome Trust Sanger Institute

Center code: SC

Web site: <http://www.sanger.ac.uk>

Contact: humquery@sanger.ac.uk

During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest, except on the rare
occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:

Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information

on the WORMPEP database can be found at

http://www.sanger.ac.uk/projects/C_elegans/wormpep RP23-109E21 is

from the RPCI-23 Mouse BAC Library

constructed by the group of Pieter de Jong.

For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: pBACe3.6

Location/Qualifiers

1..172246

/organism="Mus musculus"

/mol_type="genomic DNA"

/db_xref="taxon:10090"

/chromosome="2"

/clone="RP23-109E21"

/clone_lib="RPCI-23"

TITLE
JOURNAL

Unpublished

2 (bases 1 to 196286)

Worley, K.C.

Direct Submission

TITLE

JOURNAL

COMMENT

Query Match 77.6%; Score 19.4; DB 9; Length 172246;
Best Local Similarity 95.2%; Pred. No. 2.3e+02;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AAAAAACACACCACCTGAGC 21

|||||

Db 87236 AAAAAACACACCACCTGAGC 87256

RESULT 32
AC153907/c

LOCUS
DEFINITION

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

Mus musculus (house mouse)
Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 196286)

REFERENCE

AUTHORS
JOURNAL

COMMENT

AC153907 196286 bp DNA linear ROD 18-DEC-2004
Mus musculus BAC RP23-205P6 (Roswell Park Cancer Institute
(C57BL/6J Female) Mouse BAC Library) complete sequence.

AC153907

AC153907.1 GI:56710431

HTG.

Mus musculus (house mouse)

Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 196286)

Muzny, D., Adams, C., Agbai, II, O., Allen, C., Alsbrooks, S., Archer, P.,
Aredondo, H., Bandaranaike, D., Bangura, L., Beltran, B., Beltran, R.,
Beraducci, A., Biswal, K., Blyth, P., Bonham, H., Buhay, C., Burch, P.,
Cadoree, I., Canada, A., Cardenas, V., Carter, K., Cavazos, I.,
Chacko, J., Chahrour, M., Chavez, D., Chen, A., Chen, G., Chen, R.,
Cheng, M.-T., Chu, J., Clerc, K., Cockrell, R., Coyle, M., Cree, A.,
Curry, S., Dai, W., Davila, M.L., Davis, C., Davy-Carroll, L., De
Anda, C., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H.,
Donlin, J., McCauley, S., Dugan-Rocha, S., Dunn, A., Durbin, K., Fa, M.,
Dziuda, D., Egan, A., Escotto, M., Espinosa, V., Eugene, C., Fa, M.,
Fernandez, S., Fernando, P., Flagg, N., Forbes, L., Foster, P.,
Fowler, G., Fu, Q., Fuh, E., Garcia, A., Garcia, R., Garner, T.,
Gaskin, C., Gench, S., Ghose, S., Gill, R., Gonzalez, D.,
Gonzalez-Garay, M., Guevara, W., Holder, M., Haaland, W., Haeblerlen, K.,
Hall, B., Hamid, H., Hamilton, K., Harbes, B., Harris, R., Havlak, P.,
Hawes, A., Hawkins, E., Hayes, S., Hemphill, L., Hernandez, J.,
Hines, S., Hitchens, M., Hodgson, A., Hogues, M., Hollins, B.,
Howell, L. T., Hulyk, S., Hume, J., Imo, K., Jackson, A., Jackson, L.,
Jacob, L., Jiang, H., Johnson, B., Johnson, R., Kalafas, K., Kelly, S.,
Keys, T., Khan, Z., King, L., Kovar, C., Kowis, A., Kowis, C., Lara, F.,
Leal, S., Lee, K., Lee, S., Legall, F. I., Lemons, S., Lewis, B., Li, B.,
Li, Y., Li, Z., Linnell, M., Liu, Y., Liu, Y.-S., Liu, Y., Liyanage, D.,
London, P., Lopez, J., Lorenshew, L., Lozano, R., Luk, T., Madu, R.,
Maheshwari, M., Mahoney, C., Malloy, K., Mansouri, D., Martinez, E.,
McClelland, H., McPherson, J., Mercadao, C., Metzger, M.,
Milosavljevic, A., Minja, S., Morgan, M., Morris, S., Munidasa, M.,
Murray, D., Nazareth, L., Ngo, D., Nguyen, N., Norwig-Bastuagh, E.,
Nott, A., Nwaokwelu, O., Obregon, M., Ochi-Okorie, C., Odeh, E.,
Okwuonu, G., Okwuonu, K., Parker, D., Pasternak, S., Patel, B.,
Patel, V., Paul, H., Perez, A., Perez, L., Petrosino, J., Pham, T.,
Primus, E., Pu, J.-I., Puazo, M., Qin, X., Quinn, A., Quiroz, J.,
Rabata, D., Rachlin, E., Reigh, R., Ren, Y., Reuter, M., Richards, S.,
Rives, C., Rodriguez, P., Rojas, A., Ruiz, S.-J., Sana, M., Sanders, W.,
Santibanez, J., Santos, R., Savary, G., Scherer, S., Shen, H., Shen, Y.,
Sisson, I., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R.,
Svatek, A., Taylor, E., Taylor, T., Thomas, N., Thorn, R., Thornton, R.,
Trejos, Z., Usmani, K., Vargo, C., Verduzco, D., Villaseana, D., Virk, D.,
Volkov, A., Waldron, L., Walker, B., Wang, Q., Wang, S., Warren, J.,
Wei, X., Wheeler, D., Williams, G., Williams, R., Worley, K., Wright, R.,
Wu, J., Yakub, S., Yan, K., Yaun, Y., Yu, F., Zhang, J., Zhang, L.,
Zhang, Z., Zhou, J., Weinstock, G. and Gibbs, R.

Direct Submission

Unpublished

2 (bases 1 to 196286)

Worley, K.C.

Direct Submission

TITLE

JOURNAL

COMMENT

Submitted (18-DEC-2004) Human Genome Sequencing Center, Baylor
College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

Sequencing is completed to a minimum standard of double strand
coverage with a minimum of 2 clones and 2 reads with no ambiguities
or 2 chemistries with a minimum of 2 clones and 3 reads with no
ambiguities. If the sequence quality does not meet this standard,
it will be indicated in the annotation.

The repeat regions shown were identified using RepeatMasker by
Adrian Smit.

Sequence similarities were identified using Powerblast by Jinghui

Zhang.

Exon/Intron boundaries of identified genes were chosen if there were canonical splice junctions that maintained sequence continuity across the splice junctions.

FEATURES

source

1. 196286

/organism="Mus musculus"
/mol_type="Genomic DNA"
/db_xref="taxon:10090"
/clone="RP23-205F6"
complement(1. .78743)
/note="overlaps bases 1. .78743 of clone AC130816"
/function="clone overlap"

misc_feature

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region
complement(19096. .19306)
/rpt_family="Lx7"
19401. .19503
/rpt_family="TTCCC)n"
19508. .19535
/rpt_family="AT_rich"
19541. .19581
/rpt_family="CT-rich"
19586. .19910
/rpt_family="Lx7"
19919. .20070
/rpt_family="B3A"
20147. .20172
/rpt_family="AT_rich"
20575. .20749
/rpt_family="MLTIC"
21236. .21371
/rpt_family="MLTIC"
21927. .21978
/rpt_family="L1Md_F2"
21996. .22144
/rpt_family="L1Md_F2"
22204. .22287
/rpt_family="AT_rich"
22387. .22486
/rpt_family="L2"
complement(22774. .22852)
/rpt_family="ID2"
24586. .24874
/rpt_family="B4"
25007. .25529
/rpt_family="L1_Mus1"
complement(26134. .26398)
/rpt_family="L2"
complement(27104. .27271)
/rpt_family="MIR"
complement(27761. .27898)
/rpt_family="B1_Mur1"
complement(28154. .28228)
/rpt_family="ID4"
complement(28976. .29022)
/rpt_family="L2"
31282. .31409
/rpt_family="MTD"
31441. .31587
/rpt_family="L1MC3"
complement(31896. .31966)
/rpt_family="ID4"
32349. .32371
/rpt_family="TTTTC)n"
complement(32377. .32523)
/rpt_family="B1_Mus1"
32906. .32952
/rpt_family="(TG)n"
complement(33229. .33408)
/rpt_family="HAL1"

Query Match 77.6%; Score 19.4; DB 9; Length 196286;
Best Local Similarity 95.2%; Pred. No. 2.3e+02;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AAAAAACACCACCTGAGCT 22

DB 176325 ACAAACACCACTGAGCT 176305

RESULT 33

CR847836

LOCUS

DEFINITION

SEQUENCE, 9 unordered pieces.

ACCESSION CR847836

CR847836 205309 bp DNA linear HTG 10-JUN-2005
Danio rerio chromosome 14 clone DKEY-285U18, WORKING DRAFT
SEQUENCE, 9 unordered pieces.

* as soon as it is available and the accession number will
* be preserved.

1 38536: contig of 38536 bp in length
* 38537 38636: gap of 100 bp
* 38637 41093: contig of 2457 bp in length
* 41093 41193: gap of 100 bp
* 41194 45338: contig of 4145 bp in length
* 45338 45438: gap of 100 bp
* 45439 52680: contig of 7242 bp in length
* 52681 52780: gap of 100 bp
* 52781 119809: contig of 67029 bp in length
* 119810 130198: contig of 10289 bp in length
* 130199 130298: gap of 100 bp
* 130299 155049: contig of 24751 bp in length
* 155050 155149: gap of 100 bp
* 155150 157485: contig of 2336 bp in length
* 157486 157585: gap of 100 bp
* 157586 176949: contig of 19364 bp in length
* 176950 177049: gap of 100 bp
* 177050 194834: contig of 17785 bp in length
* 194835 194935: gap of 100 bp
* 194935 214525: contig of 19591 bp in length.

FEATURES

Source

1..214525
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="DKEY-151K14"
/clone_lib="DanioKey"

misc_feature

1..38536
/note="assembly_fragment:01440
fragment_chain:1"

misc_feature

38637..41093
/note="assembly_fragment:00014
fragment_chain:1"

misc_feature

41194..45338
/note="assembly_fragment:00046
fragment_chain:1"

misc_feature

45439..52680
/note="assembly_fragment:00080
fragment_chain:2"

misc_feature

52781..119809
/note="assembly_fragment:02130
fragment_chain:2"

misc_feature

119910..130198
/note="assembly_fragment:00115
fragment_chain:3"

misc_feature

130299..155049
/note="assembly_fragment:00540
fragment_chain:3"

misc_feature

155150..157485
/note="assembly_fragment:00001"
fragment_chain:3"

misc_feature

157586..176949
/note="assembly_fragment:00245.0"
fragment_chain:3"

misc_feature

177050..194834
/note="assembly_fragment:00878.0"
fragment_chain:3"

misc_feature

194935..214525
/note="assembly_fragment:00878.1"

ORIGIN

Query Match 77.6%; Score 19.4; DB 14; Length 214525;
Best Local Similarity 95.2%; Pred. No. 2.3e+02;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 AACAAACCACCTGAGCTGGG 25

Db 162320 AACAAACCACCTGAGCTGTG 162340

RESULT 35

AC105669

LOCUS

261112 bp DNA linear

HTG 08-OCT-2002

DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

Rattus norvegicus clone CH230-64A8, *** SEQUENCING IN PROGRESS ***.
AC105669
AC105669.3 GI:22857190
HTG: HTGS_PHASE2; HTGS_DRAFT; HTGS_ENRICHED.
Rattus norvegicus (Norway rat)
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Rattus.

REFERENCE

AUTHORS

1 (bases 1 to 261112)
Muzny, D. Marie., Metzker, M. Lee., Abramzon, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,
Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G.,
Fernandez, S., Finley, M., Flag, N., Forbes, L., Foster, M., Foster, P.,
Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
Gebregorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M.,
Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
Karpathy, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
Lorensu, H., Louised, H., Lozada, R. J., Lu, X., Ma, J.,
Maheshwari, M., Mahindartne, M., Mahmoud, M., Mallory, K., Mangum, A.,
Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
Mawhiney, S., McLeod, M. P., McNeill, T. Z., Meenen, E.,
Milosavljevic, A., Minet, G., Minja, E., Montemayor, J., Moore, S.,
Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Naif, L.,
Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
Nwackeleme, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K.,
Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,
Plommer, F., Poindexter, A., Popovic, D., Primus, E., Pu, L. L.,
Puzo, M., Quiroz, J., Rachin, E., Reeves, K., Regier, M. A., Reigh, R.,
Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J.,
Sanders, W., Savery, G., Scherer, S., Scott, G., Shatman, S., Shen, H.,
Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smajs, D.,
Sneed, A., Sodergren, E., Song, X. Z., Sorelle, R., Sosa, J.,
Steimle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C.,
Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Umanli, K.,
Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J.,
Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
Williams, G., Willson, R., Wleczyk, R., Wooden, H., Worley, K.,
Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O.,
Weinstock, G. and Gibbs, R. A.

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

On Sep 14, 2002 this sequence version replaced gi:21736569.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GNMF
Center clone name: CH230-64A8
----- Summary Statistics

Assembly program: Phrap; version 0.990329
Consensus quality: 229400 bases at least Q40
Consensus quality: 232950 bases at least Q30
Consensus quality: 235263 bases at least Q20
Estimated insert size: 253424; sum-of-contigs estimation
Quality coverage: 4x in Q20 bases; sum-of-contigs estimation

- * NOTE: Estimated insert size may differ from sequence length (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)
- * NOTE: This sequence may represent more than one clone.
- * NOTE: This is a 'working draft' sequence. It currently consists of 1 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.
- * This sequence will be replaced
- * by the finished sequence as soon as it is available and the accession number will be preserved.
- * 1 261112: contig of 261112 bp in length.

FEATURES

source
1. .261112
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-64A8"
1. .1063
/note="wgs contig"
3350. .5003
/note="wgs contig"
136644. .137706
/note="wgs_contig"

ORIGIN

Query Match 77.6%; Score 19.4; DB 14; Length 261112;
Best Local Similarity 95.2%; Pred. No. 2.3e+02;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 AAAAACACACCACTGAGCT 22
Db 38301 AAAAAAACCACTGAGCT 38321

RESULT 36

AF302834
LOCUS AF302834 836 bp DNA linear INV 17-DEC-2000
DEFINITION Haliotis rubra clone 2.29 microsatellite VNTR sequence.
ACCESSION AF302834
VERSION AF302834.1 GI:11875757
KEYWORDS
SOURCE Haliotis rubra

ORGANISM

Haliotis rubra
Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda;
Vetigastropoda; Haliotoidea; Haliotidae; Haliotis.

REFERENCE

1 (bases 1 to 836)
Evans,B.S.
Microsatellite marker conservation in the genus Haliotis

AUTHORS

Unpublished

JOURNAL

2 (bases 1 to 836)

REFERENCE

Evans,B.S.

Direct Submission

TITLE

Submitted (06-SEP-2000) Aquaculture and Biotechnology, CSIRO Marine

JOURNAL

Research, GPO Box 1538, Hobart, Tasmania 7001, Australia

FEATURES

Location/Qualifiers

1. .836
/organism="Haliotis rubra"

/mol_type="genomic DNA"

/db_xref="taxon:36100"

/clone="2.29"

218. .238

322. .437

/note="microsatellite VNTR"

/rpt_type=tandem

complement(519. .538)

primer_bind

ORIGIN

Query Match 76.8%; Score 19.2; DB 2; Length 836;

Best Local Similarity 87.5%; Pred. No. 2.6e+02;

Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AAAAAACACCACTGAGCTG 24

Db 633 AAAAAACACCACTGAGCTG 656

RESULT 37

AL627225/c

LOCUS

DEFINITION

Human DNA sequence from clone RP11-100N5 on chromosome 9 Contains part of the AKAP2 (PALM2) gene for A kinase (PRKA) anchor protein 2 (paralemmin 2), complete sequence.

AL627225

AL627225.11 GI:17066061

HTG; AKAP2; anchor protein; CpG island; PALM2; paralemmin.

KEYWORDS

SOURCE

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

1 (bases 1 to 3994)

Pandian,R.

Direct Submission

Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,

Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vega@sanger.ac.uk

Clone requests: clonerequest@sanger.ac.uk

On Nov 25, 2001 this sequence version replaced gi:17017834.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:

Em., EMBL; Sw., SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information

in the WORMPEP database can be found at

http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence

was generated from part of bacterial clone contigs of human

chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping

Group. Further information can be found at

<http://www.sanger.ac.uk/HGP/Chr9>

RP11-100N5 is from the library RPI-11.1 constructed by the group

of Pieter de Jong. For further details see

<http://www.chori.org/bacpac/home.htm>

VECTOR: pBACe3.6

----- Genome Center

Center: Wellcome Trust Sanger Institute

Center code: SC

Web site: <http://www.sanger.ac.uk>

Contact: vega@sanger.ac.uk


```

TNAQENSLADFLSPOTPDNPSEGRGEGVSKSFSDHGFYSPSSTLGDSPVLDDPLE
YQAGLLVONAIQOAIQAEVDKAVSKTRSDGAEOQPEATVEBAEAAAFSGEKPSQMF
PROVSPVOEKEDVLKILPAEDRALRERGPQPLPAVOPSPGINWEETRPGSVFSK
YSEAEIRSTASLATQESDVMVCPKLSRKQRTLSMEERIRAQEREELKQRQ
VLOSTQSPRTKNAPKPTCYKTAFGKJEKVPPSPPTGSLQDFDAPBEAAGTQ
RPKNLQMTLMEDYETHKSKRRERMDSSYTKLLSCKVTSEVLEATRVNRRKSALALR
WEAGIYANOEEEDNE"
join(AL158829,14:38598..38642,AL158829,14:125600..125680,
3937..3994,AL353806,9:1999..2071,AL353806,9:45125..45239,
AL353806,9:46451..46472,AL353806,9:53317..53418,
AL353806,9:64076..64815)
/gene="AKAP2"
/locus tag="RP11-470J20.3-001"
/standard name="OTTHUMP0000021890"
/notes="match: proteins: Q96DUI"
/codon_start=1
/product="A kinase (PRKA) anchor protein 2"

Query Match      76.8%; Score 19.2; DB 8; Length 3994;
Best Local Similarity 87.5%; Pred. No. 2.7e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAAACACCACTGAGCTGG 25
    ||||| ||||| ||||| |||||
Db 2439 AAAAAAACCACTGGCTGG 2416

RESULT 38
AC104180/c
LOCUS
DEFINITION
AC104180
VERSION
AC104180.5 GI:29244611
KEYWORDS
HTG; HTGS PHASE2.
SOURCE
Leishmania major
ORGANISM
Eukaryota; Euzlenozoa; Kinetoplastida; Trypanosomatidae;
Leishmania.
REFERENCE
1 (bases 1 to 90100)
Myler,P.J., Sisk,E., Nelson,S., Mack,J., Vogt,C., Robertson,L.,
Worthey,E., Ivens,A., Marty,A., Munden,H., Stuart,K., Seyler,A. and
Rinta,J.
Direct Submission
Submitted (06-DEC-2001) Seattle Biomedical Research Institution, 4
Nickerson Street, Seattle, WA 98109-1651, USA
2 (bases 1 to 90100)
Myler,P.J., Sisk,E., Vogt,C., Robertson,L., Mack,J., Nelson,S.,
Worthey,E., Ivens,A., Marty,A., Munden,H., Stuart,K., Seyler,A. and
Rinta,J.
Direct Submission
Submitted (26-MAR-2003) Seattle Biomedical Research Institution, 4
Nickerson Street, Seattle, WA 98109-1651, USA
On Mar 26, 2003 this sequence version replaced gi:25447632.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 68517: contig of 68517 bp in length
* 68518 68650: gap of unknown length
* 68651 90100: contig of 21450 bp in length.
FEATURES
Location/Qualifiers
1..90100
/organism="Leishmania major"
/mol_type="genomic DNA"
/strain="Friedlin"
/db_xref="taxon:5664"
/chromosome="22"

/clone="LB00273"
68518..68650
/estimated_length=unknown

gap

ORIGIN

Query Match      76.8%; Score 19.2; DB 14; Length 90100;
Best Local Similarity 87.5%; Pred. No. 2.8e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAACACCACTGAGCTGG 24
    ||||| ||||| ||||| |||||
Db 15628 AAAAAAGAACCACTGGAGCGG 15605

RESULT 39
AL359978,0/c
WPCOMMENT
Sequence split into 4 fragments LOCUS AL359978 Accession AL359978
Fragment Name      Begin      End
AL359978.0         1      110000
AL359978.1        100001    210000
AL359978.2        200001    310000
AL359978.3        300001    387346
LOCUS AL359978
DEFINITION Homo sapiens chromosome 9 clone RP11-144G15, 54 unordered pieces.
ACCESSION AL359978
VERSION AL359978.16 GI:14575237
KEYWORDS HTG; HTGS PHASE1; HTGS_CANCELLED.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
1
Plumb,B.
Direct Submission
Submitted (14-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Jun 28, 2001 this sequence version replaced gi:12331081.
----- Genome Center
Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: BA144G15
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: plasmid; L08752; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 364536 bases at least Q40
Consensus quality: 372753 bases at least Q30
Consensus quality: 377828 bases at least Q20
Insert size: 382046; sum-of-contigs
Insert size: 172123; 13.2% error; agarose-fp
Quality coverage: 8.09x in Q20 bases; sum-of-contigs Quality
coverage: 22.67x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 54 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 35539: contig of 35539 bp in length
* 35540 35639: gap of 100 bp
* 35640 46944: contig of 11305 bp in length
* 46945 47044: gap of 100 bp
* 47045 66901: contig of 19857 bp in length
* 66902 67001: gap of 100 bp
```

* 67002	70175: contig of 3174 bp in length	
* 70176	70275: gap of 100 bp	
* 70276	104443: contig of 34168 bp in length	
* 104444	104543: gap of 100 bp	
* 104544	181879: contig of 77336 bp in length	
* 181880	181979: gap of 100 bp	
* 181980	195035: contig of 13056 bp in length	
* 195036	195135: gap of 100 bp	
* 195136	197887: contig of 2752 bp in length	
* 197888	197987: gap of 100 bp	
* 197988	200370: contig of 2383 bp in length	
* 200371	200470: gap of 100 bp	
* 200471	204730: contig of 4260 bp in length	
* 204731	204830: gap of 100 bp	
* 204831	211277: contig of 6447 bp in length	
* 211278	211377: gap of 100 bp	
* 211378	215636: contig of 4259 bp in length	
* 215637	215736: gap of 100 bp	
* 215737	219304: contig of 3568 bp in length	
* 219305	219404: gap of 100 bp	
* 219405	221607: contig of 2203 bp in length	
* 221608	221707: gap of 100 bp	
* 221708	226929: contig of 5222 bp in length	
* 226930	227029: gap of 100 bp	
* 227030	231643: contig of 4614 bp in length	
* 231644	231743: gap of 100 bp	
* 231744	235404: contig of 3661 bp in length	
* 235405	235504: gap of 100 bp	
* 235505	246223: contig of 10721 bp in length	
* 246226	246325: gap of 100 bp	
* 246326	253367: contig of 7042 bp in length	
* 253368	253467: gap of 100 bp	
* 253468	258279: contig of 4812 bp in length	
* 258280	258379: gap of 100 bp	
* 258380	261753: contig of 3374 bp in length	
* 261754	261853: gap of 100 bp	
* 261854	264531: contig of 2678 bp in length	
* 264532	264631: gap of 100 bp	
* 264632	270234: contig of 5603 bp in length	
* 270235	270334: gap of 100 bp	
* 270335	273988: contig of 3654 bp in length	
* 273989	274088: gap of 100 bp	
* 274089	277769: contig of 3681 bp in length	
* 277770	277869: gap of 100 bp	
* 277870	283666: contig of 5797 bp in length	
* 283667	283766: gap of 100 bp	
* 283767	286170: contig of 2404 bp in length	
* 286171	286270: gap of 100 bp	
* 286271	288379: contig of 2109 bp in length	
* 288380	288479: gap of 100 bp	
* 288480	291608: contig of 3129 bp in length	
* 291609	291708: gap of 100 bp	
* 291709	296427: contig of 4719 bp in length	
* 296428	296527: gap of 100 bp	
* 296528	302497: contig of 5970 bp in length	
* 302498	302597: gap of 100 bp	
* 302598	319723: contig of 17126 bp in length	
* 319724	319823: gap of 100 bp	
* 319824	322858: contig of 3035 bp in length	
* 322859	322958: gap of 100 bp	
* 322959	325608: contig of 2650 bp in length	
* 325609	325708: gap of 100 bp	
* 325709	329450: contig of 3742 bp in length	
* 329451	329550: gap of 100 bp	
* 329551	331984: contig of 2434 bp in length	
* 331985	332084: gap of 100 bp	
* 332085	334354: contig of 2270 bp in length	
* 334355	334454: gap of 100 bp	
* 334455	337261: contig of 2807 bp in length	
* 337262	337361: gap of 100 bp	
* 337362	339476: contig of 2115 bp in length	
* 339477	339576: gap of 100 bp	
* 339577	342596: contig of 3020 bp in length	

* 342597	342696: gap of 100 bp	
* 342697	345524: contig of 2828 bp in length	
* 345525	345624: gap of 100 bp	
* 345625	347670: contig of 2046 bp in length	
* 347671	347770: gap of 100 bp	
* 347771	352927: contig of 5157 bp in length	
* 352928	353027: gap of 100 bp	
* 353028	355067: contig of 2040 bp in length	
* 355068	355167: gap of 100 bp	
* 355168	357536: contig of 2369 bp in length	
* 357537	357636: gap of 100 bp	
* 357637	365237: contig of 7601 bp in length	
* 365238	365337: gap of 100 bp	
* 365338	367721: contig of 2384 bp in length	
* 367722	367821: gap of 100 bp	
* 367822	370537: contig of 2716 bp in length	
* 370538	370637: gap of 100 bp	
* 370638	374038: contig of 3401 bp in length	
* 374039	374138: gap of 100 bp	
* 374139	376389: contig of 2251 bp in length	
* 376390	376489: gap of 100 bp	
* 376490	379106: contig of 2617 bp in length	
* 379107	379206: gap of 100 bp	
* 379207	382405: contig of 3199 bp in length	
* 382406	382505: gap of 100 bp	
* 382506	384798: contig of 2293 bp in length	
* 384799	384898: gap of 100 bp	
* 384899	387346: contig of 2448 bp in length.	

FEATURES	Location/Qualifiers	
source	1..387346	
	/organism="Homo sapiens"	
	/mol_type="genomic DNA"	
	/db_xref="taxon:9606"	
	/chromosome="9"	
	/clone="RP11-144G15"	
	/clone_lib="RPC1-11.1"	
misc_feature	1..35539	
	/note="assembly fragment:05228	
	fragment_chain:1	
misc_feature	35640..46944	
	/note="assembly fragment:09756	
	fragment_chain:1	
misc_feature	47045..66901	
	/note="assembly fragment:00410	
	fragment_chain:1	
misc_feature	67002..70175	
	/note="assembly fragment:01425	
	fragment_chain:1	
misc_feature	70276..104443	
	/note="assembly fragment:09780	
	fragment_chain:1	
misc_feature	104544..181879	
	/note="assembly fragment:05688	
	fragment_chain:1	
misc_feature	181980..195035	
	/note="assembly fragment:03194	
	fragment_chain:1	
misc_feature	195136..197887	
	/note="assembly fragment:02620	
	fragment_chain:2	
misc_feature	197988..200370	
	/note="assembly fragment:03650	
	fragment_chain:2	
misc_feature	200471..204730	
	/note="assembly fragment:00487	
	fragment_chain:2	
misc_feature	204831..211277	
	/note="assembly fragment:04780	
	fragment_chain:2	
misc_feature	211378..215636	

Query Match

Best Local Similarity

76.8%; Score 19.2; DB 14; Length 110000;

87.5%; Pred. No. 2.8e+02;

Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;		
Qy	2 AAAAAACCAACCACTGAGCTGG 25 	
Db	104982 AAAAAACCAACCACTGGCTGG 104959	
RESULT 40		
AL359978 1/c		
WPCOMMENT		
Sequence split into 4 fragments LOCUS AL359978 Accession AL359978		
Fragment Name	Begin	End
AL359978_0	1	110000
AL359978_1	100001	210000
AL359978_2	200001	310000
AL359978_3	300001	387346
Continuation (2 of 4) of AL359978 from base 100001 (AL359978 Homo sapiens chromosome 9 c		
Query Match 76.8%; Score 19.2; DB 14; Length 110000;		
Best Local Similarity 87.5%; Pred. No. 2.8e+02;		
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;		
Qy	2 AAAAAACCAACCACTGAGCTGG 25 	
Db	4982 AAAAAACCAACCACTGGCTGG 4959	
RESULT 41		
BX936368 2/c		
WPCOMMENT		
Sequence split into 5 fragments LOCUS BX936368 Accession BX936368		
Fragment Name	Begin	End
BX936368_0	1	110000
BX936368_1	100001	210000
BX936368_2	200001	310000
BX936368_3	300001	410000
BX936368_4	400001	429582
Continuation (3 of 5) of BX936368 from base 200001 (BX936368 Danio rerio clone DKEY-1891		
Query Match 76.8%; Score 19.2; DB 14; Length 110000;		
Best Local Similarity 87.5%; Pred. No. 2.8e+02;		
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;		
Qy	2 AAAAAACCAACCACTGAGCTGG 25 	
Db	71918 AAAAAACCAACCCCTGTTCTGG 71895	
RESULT 42		
CT005261_0		
WPCOMMENT		
Sequence split into 8 fragments LOCUS CT005261 Accession CT005261		
Fragment Name	Begin	End
CT005261_0	1	110000
CT005261_1	100001	210000
CT005261_2	200001	310000
CT005261_3	300001	410000
CT005261_4	400001	510000
CT005261_5	500001	610000
CT005261_6	600001	710000
CT005261_7	700001	716608
LOCUS CT005261 716608 bp DNA linear HTG 04-JUL-2005		
DEFINITION Leishmania major strain Friedlin, *** SEQUENCING IN PROGRESS ***.		
ACCESSION CT005261 AC008240		
VERSION CT005261.1 GI:68224038		
KEYWORDS HTG; HTGS PHASE2.		
SOURCE Leishmania major		
ORGANISM Leishmania major		
Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;		
Leishmania.		
1 (bases 1 to 716608)		
REFERENCE Peacock,C.S., Murphy,L., Ivens,A.C., Berriman,M., Blackwell,J.,		
AUTHORS Smith,D., Collins,M., Fosker,N., Harris,D., Oliver,K., O'Neil,S.,		

Saunders,D., Seeger,K., Warren,T., Rajandream,M. and Barrell,B.G.		
Direct Submission		
Submitted (15-JUN-2005) Peacock C.S., The Wellcome Trust Sanger		
Institute Genome Campus, Hinxton, Cambridge CB10 1SA, UK		
On Jul 6, 2005 this sequence version replaced gi:29825871.		
* NOTE: This is a 'working draft' sequence. It currently		
* consists of 1 contigs. Gaps between the contigs		
* are represented as runs of N. The order of the pieces		
* is believed to be correct as given, however the sizes		
* of the gaps between them are based on estimates that have		
* provided by the submitter.		
* This sequence will be replaced		
* by the finished sequence as soon as it is available and		
* the accession number will be preserved.		
* 1 716608: contig of 716608 bp in length.		
* Location/Qualifiers		
FEATURES	source	
1..716608		
/organism="Leishmania major"		
/mol_type="genomic DNA"		
/strain="Friedlin"		
/db_xref="taxon:5664"		
/note="LmjF22_01.20050601_V5.2"		
complement(3313..5241)		
/locus_tag="LmjF22.0010"		
complement(3313..5241)		
/locus_tag="LmjF22.0010"		
/note="LmjF22.0010, predicted protein, len = 643 aa,		
unknown protein; predicted pi = 8.0047"		
/codon_start=1		
/product="hypothetical protein, conserved"		
/protein_id="CAJ03760.1"		
/db_xref="GI:68224039"		
/translation="MRGVDKGPGREGVEVDAGVVTEDLRKDGVDNPSISLSON		
EPGIEESDRYKRDATVHLKRTIIRNSVFLVLGVFNPHYCLIMSAASLAEGY		
GLKLVALLITWNIIPGIFARVLNMFATVSYNVRITAMSLSLGIRFVSFAAPIG		
GYNNSAFVVMGLTGVLLGTASTYGSVMLTFVQRPDSIVGAMSGTGISGVAASLI		
FGLTSAGTQQOTFLVITPLCIITWLCFMLGNVSPRKVLATVTRNGRDIDILCEP		
GVEDAHICLLHRAARHVDVSEDTIKVMNIRAVSHQSPVHVHPSPSABEGDG		
LVYPCRVTCGCCSPNGRWRWDVGPDLIFMNTMAWFFNLAVVYAYAAQLMA		
PFSFYCEPEWKNFVKNVYVVCQYQLGVLISRSLLIKVPIYGVVISIIQVNAV		
CHIVQARVMYIKTSRQVGLSFILFAMWIFGLMGASVNVVLILKRSITLRR		
EREAEVAYLTSKARDVSGRRYADELRKKAADERAALVAKLEKPVSEIMPKPSQ		
SDTASAPQSQILGLATQQRDAIKAIHCLNDVWAAREMGMINIGALYATVGTITLG		
TVVDLIFTNTMLKHSKSC"		
complement(join(3349..3408,3751..3819,3862..3930,		
3949..4008,4051..4119,4480..4548,4576..4644,4663..4731,		
4759..4827,4864..4923,4951..5019))		
/locus_tag="LmjF22.0010"		
/note="11 probable transmembrane helices predicted for		
LmjF22.0010 by TMHMM2.0 at aa 75-97, 107-126, 139-161,		
171-193, 200-222, 232-254, 375-397, 412-431, 438-460,		
475-497 and 612-631"		
complement(3676..5025)		
/locus_tag="LmjF22.0010"		
/note="HMMPFam		
HMMPFam hit to PF02487, CLN3 protein, score 1.7e-08"		
complement(11185..11619)		
/locus_tag="LmjF22.0020"		
complement(11185..11619)		
/locus_tag="LmjF22.0020"		
/note="LmjF22.0020, predicted protein, len = 643 aa,		
unknown; predicted pi = 5.5753"		
/codon_start=1		
/product="hypothetical protein, unknown function"		
/protein_id="CAJ03762.1"		
/db_xref="GI:68224040"		
/translation="MSSTVPHAAASAKVVDCKEIGVTIDLHSPPTASNFLAYNRVCSA		
LDSDYLMQPINEAQYDSETOACVDLSLNGPTSFYDHYVRKPTRKRLGVLAVGESPMEHV		
GVSEDFLMPAQFDLPAAACAQACAFSDPLTRLATGRWPPRD"		
complement(12239..12805)		
/locus_tag="LmjF22.0030"		
complement(12239..12805)		
/locus_tag="LmjF22.0030"		

REFERENCE 4 (bases 1 to 134793)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Submitted (17-JAN-2003) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT On Jan 17, 2003 this sequence version replaced gi:22038504.
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www.shgc.stanford.edu
Quality: Phrap Quality >=40 99.4% of Sequence;
Estimated Total Number of Errors is 0.4.
NOTE: Shatter libraries failed to resolve dinucleotide repeat region 104785-104880. Unsure number of repeat copies 104785-104880. Forced join 104826.
FEATURES
source 1. .134793
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="19"
/clone="CTD-2105E13"
104785..104880
misc_feature
/notes="NOTE: Shatter libraries failed to resolve dinucleotide repeat region 104785-104880. Unsure number of repeat copies 104785-104880. Forced join 104826."
ORIGIN
Query Match 76.8%; Score 19.2; DB 8; Length 134793;
Best Local Similarity 87.5%; Pred. No. 2.9e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 2 AAAAACAACCACTGAGCTGGG 25
Db 59121 AAATACAACGACGACTGACCTGGG 59144
RESULT 47
AC154498 138861 bp DNA linear HTG 30-DEC-2004
LOCUS Mus musculus chromosome 17 clone RP24-112F5, WORKING DRAFT
DEFINITION
SEQUENCE, 7 unordered pieces.
AC154498 AC107729
AC154498.1 GI:56900124
HTG; HTGS PHASE1; HTGS DRAFT; HTGS ACTIVEFIN.
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Muridae; Muridae; Mus.
Wilson, R.K.
1 (bases 1 to 138861)
TITLE The sequence of Mus musculus clone
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 138861)
Wilson, R.K.
Direct Submission
TITLE Submitted (30-DEC-2004) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
JOURNAL
COMMENT On Dec 30, 2004 this sequence version replaced gi:39841434.
----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu
Contact: submissions@genome.wustl.edu
----- Project Information -----
Center project name: M_BB0112F05
Drafting center: WIBR
----- Summary Statistics -----
Sequencing vector: M13; 0%
Sequencing vector: plasmid; 100%

Chemistry: Dye-primer ET; 0% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 135784 bases at least Q40
Consensus quality: 136832 bases at least Q30
Consensus quality: 137476 bases at least Q20

* NOTE: This is a 'working draft' sequence. It currently consists of 7 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. * This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.
* 1 1379: contig of 1379 bp in length
* 1380 1479: gap of unknown length
* 1480 2674: contig of 1195 bp in length
* 2675 2774: gap of unknown length
* 2775 3800: contig of 1026 bp in length
* 3801 3900: gap of unknown length
* 3901 12062: contig of 8162 bp in length
* 12063 12162: gap of unknown length
* 12163 26951: contig of 14789 bp in length
* 26952 27051: gap of unknown length
* 27052 70763: contig of 43712 bp in length
* 70764 70863: gap of unknown length
* 70864 138861: contig of 67998 bp in length.
FEATURES
Location/Qualifiers
source 1. .138861
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="17"
/clone="RP24-112F5"
1. .1379
misc_feature
/notes="assembly_name:Contig86"
gap 1380..1479
/estimated_length=unknown
misc_feature 1480..2674
/notes="assembly_name:Contig89"
gap 2675..2774
/estimated_length=unknown
misc_feature 2775..3800
/note="assembly_name:Contig90"
gap 3801..3900
/estimated_length=unknown
misc_feature 3901..12062
/notes="assembly_name:Contig91"
gap 12063..12162
/estimated_length=unknown
misc_feature 12163..26951
/notes="assembly_name:Contig92"
gap 26952..27051
/estimated_length=unknown
misc_feature 27052..70763
/notes="assembly_name:Contig93"
gap 70764..70863
/estimated_length=unknown
misc_feature 70864..138861
/note="assembly_name:Contig94"
ORIGIN
Query Match 76.8%; Score 19.2; DB 14; Length 138861;
Best Local Similarity 87.5%; Pred. No. 2.9e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 2 AAAAACAACCACTGAGCTGGG 25
Db 107216 AAAAACAACCACTGAGCTGGG 107239
RESULT 48

```

AC145018/c
LOCUS       AC145018               141663 bp    DNA    linear    HTG 26-JUN-2003
DEFINITION   Felis catus clone RP86-279N10, WORKING DRAFT SEQUENCE, 5 ordered
             pieces.
ACCESSION   AC145018
VERSION     AC145018.2   GI:32261343
KEYWORDS    HTG; HTGS_PHASE2; HTGS_DRAFT.
SOURCE      Felis catus (cat)
ORGANISM    Felis catus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Felidae;
            Felinae; Felis.
REFERENCE   1 (bases 1 to 141663)
AUTHORS    Antonellis A., Ayele K., Beckstrom-Sternberg, S.M., Benjamin B.,
            Blakesley R.W., Bouffard G.G., Brinkley C., Brooks S., Cariaga K.,
            Chu G., Coleman B., Coleman H., Engle J., Granice S., Guan X.,
            Gupta J., Haghighi P., Han J., Hansen N., Ho S.-L., Hu P.,
            Hurle B., Idol J.R., Karlins E., Kwong P., Laric P., Lee-Lin S.-Q.,
            Legaspi R., Maduro Q.L., Maduro V.B., Margulies E.H., Masiello C.,
            Maskeri B., McDowell J., Paguirigan C., Pearson R., Portnoy M.E.,
            Prasad A., Reddix-Dugue N., Schandler K., Schueler M.G., Shah K.,
            Sison C., Stantripop S., Thomas J.W., Thomas P.J., Tsipouri V.,
            Vogt J.L., Wetherby K.D., Wiggins L., Young A. and Green E.D.
            NISC Comparative Sequencing Initiative
            Unpublished
            2 (bases 1 to 141663)
            Green, E.D.
            Direct Submission
            Submitted (30-MAY-2003) NIH Intramural Sequencing Center, 8717
            Grovemont Circle, Gaithersburg, MD 20877, USA
            3 (bases 1 to 141663)
            Green, E.D.
            Direct Submission
            Submitted (26-JUN-2003) NIH Intramural Sequencing Center, 8717
            Grovemont Circle, Gaithersburg, MD 20877, USA
            On Jun 26, 2003 this sequence version replaced gi:31193896.
            ----- Genome Center
            Center: NIH Intramural Sequencing Center
            Center code: NISC
            Web site: http://www.nisc.nih.gov
            Contact: nisc.zoo@nhgri.nih.gov
            ----- Project Information
            Center project name: efl
            Center clone name: 279N10

The sequence data in this record represents an 'enhanced'
version of a Phase 2 submission. Specifically, the indicated
order and orientation of each sequence contig has been
established using one or more of the following: read-pair
data from individual subclones, overlaps with neighboring
clones, alignment with available reference sequence (e.g.,
human), and/or confirmation by PCR testing. In addition,
the sequence assembly is based on at least 8X average
coverage in Q20 bases and has been reviewed to rule out
gross misassemblies, the low-quality ends of sequence
contigs have been trimmed away, and each base is associated
with a Phrap-derived quality score.
----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 140709 bases at least Q40
Consensus quality: 140932 bases at least Q30
Consensus quality: 141174 bases at least Q20
Insert size: 123000; agarose-fp
Insert size: 141263; sum-of-contigs
Quality coverage: 11.81x in Q20 bases; agarose-fp
Quality coverage: 10.28x in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 5 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes

```

* of the gaps between them are based on estimates that have
* provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 755: contig of 755 bp in length
* 756 855: gap of unknown length
* 856 8935: contig of 8080 bp in length
* 8936 9035: gap of unknown length
* 9036 32437: contig of 23401 bp in length
* 32437 32537: gap of unknown length
* 32537 55738: contig of 23201 bp in length
* 55738 55837: gap of unknown length
* 55837 141663: contig of 85826 bp in length.
* Location/Qualifiers
* 1..141663
* /organism="Felis catus"
* /mol_type="genomic DNA"
* /db_xref="taxon:9685"
* /clone="RP86-279N10"
* /clone_lib="RP86"
* 1..755
* /note="assembly_fragment
* clone_end:T7
* vector_side:left"
* 756..855
* /estimated_length=unknown
* 856..8935
* /note="assembly_fragment"
* 8936..9035
* /estimated_length=unknown
* 9036..32436
* /note="assembly_fragment"
* 32437..32536
* /estimated_length=unknown
* 32537..55737
* /note="assembly_fragment"
* 55738..55837
* /estimated_length=unknown
* 55838..141663
* /note="assembly_fragment
* clone_end:SP6
* vector_side:right"

ORIGIN

Query Match 76.8%; Score 19.2; DB 14; Length 141663;
Best Local Similarity 87.5%; Pred.No. 2.9e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAAACACCCACTGAGCTGGG 25
|||||
Db 124266 AAAAAACACCCACTGAGCAGGG 124243

RESULT 49
AC157291
LOCUS AC157291 Bos taurus clone CH240-72B6, WORKING DRAFT SEQUENCE, 7 unordered
DEFINITION pieces.
ACCESSION AC157291
VERSION AC157291.2 GI:60301583
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
Pecora; Bovidae; Bovinae; Bos.
1 (bases 1 to 154763)
Muzny, D. Marie., Metzker, M. Lee., Abramson, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
Anyalebechi, V., Aoyagi, A., Ayodeji, I. M., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,

Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, P., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escoto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabis, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebrgeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hulvy, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, H., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpach, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensuhewa, L., Loulseghe, H., Lozada, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwaokeme, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plapper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L., Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajls, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villaseña, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wlezyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiser, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

Direct Submission
Unpublished
2 (bases 1 to 154763)
Worley, K.C.
Direct Submission
Submitted (11-FEB-2005) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 154763)
Cow Genome Sequencing Consortium.
Direct Submission
Submitted (01-JUL-2005) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Jun 29, 2005 this sequence version replaced gi:59276096.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

Center: Baylor College of Medicine

Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
Project Information
Center project name: FGCN
Center clone name: CH240-7286
Assembly program: Atlas 3.0;
Consensus quality: 152641 bases at least Q40
Consensus quality: 153258 bases at least Q30
Consensus quality: 153807 bases at least Q20
Estimated insert size: 156808; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

NOTE: Estimated insert size may differ from sequence length (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
NOTE: This is a 'working draft' sequence. It currently consists of 7 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 2299: contig of 2299 bp in length
2300: gap of 83 bp
2382: contig of 9618 bp in length
12001: gap of 50 bp
12050: gap of 26562 bp in length
12051: gap of 153 bp
38765: gap of 16952 bp in length
38766: gap of 50 bp
55767: gap of 50 bp
99117: contig of 43350 bp in length
99157: gap of 50 bp
99168: gap of 50 bp
99168: gap of 50 bp
144531: contig of 45363 bp in length
144531: gap of 50 bp
144581: contig of 10183 bp in length.

FEATURES

source
/organism="Bos taurus"
/mol_type="genomic DNA"
/db_xref="taxon:9913"
/clone="CH240-7286"
2300..2382
/estimated_length=83
12001..12050
/estimated_length=50
38613..38765
/estimated_length=153
55718..55767
/estimated_length=50
99118..99167
/estimated_length=50
144531..144580
/estimated_length=50

ORIGIN

Query Match 76.8%; Score 19.2; DB 14; Length 154763;
Best Local Similarity 87.5%; Pred. No. 2.9e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 2 AAAAAACACCACTGAGCTGGG 25
|||||
Db 147405 AAAAAACACCACTGAGCTGGG 147428
|||||
RESULT 50
AC016868/c
LOCUS AC016868 155691 bp DNA linear PRI 06-FEB-2002
DEFINITION Homo sapiens chromosome 8, clone RP11-15G16, complete sequence.
ACCESSION AC016868
VERSION AC016868.8 GI:18377299
KEYWORDS HTG.


```
repeat_region /rpt_family="L1ME4A"
17327..17569
/rpt_family="MER106A"
complement(17570..17793)
/rpt_family="L1ME4A"
complement(17794..18102)
/rpt_family="AluSx"
complement(18103..18407)
/rpt_family="L1ME4A"
complement(18409..18588)
/rpt_family="AluSg/x"
18589..18842
/rpt_family="AluSx"
complement(19365..19666)
/rpt_family="AluSx"
complement(19730..20148)
/rpt_family="L1ME3"
complement(20173..20384)
/rpt_family="MLTID"
complement(20445..20505)
/rpt_family="MLTID"
complement(20506..20804)
/rpt_family="AluSx"
21254..21280
/rpt_family="(CA)n"
21282..21314
/rpt_family="(TA)n"
complement(21320..21471)
/rpt_family="L1ME2"
complement(21472..21635)
/rpt_family="FRAM"
complement(21636..22021)
/rpt_family="L1ME2"
complement(22022..22063)
/rpt_family="THE1C"
22064..22374
/rpt_family="AluSx"
complement(22312..22316)
/note="<30 qual SNGL region"
complement(22337..22346)
/note="<30 qual SNGL region"
complement(22375..22746)
/rpt_family="THE1C"
complement(22747..22999)
/rpt_family="L1ME2"
complement(23000..23306)
/rpt_family="AluSq"
complement(23307..24065)
/rpt_family="L1ME2"
```

Query Match 76.8%; Score 19.2; DB 8; Length 155691;
Best Local Similarity 87.5%; Pred.No. 2.9e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AAAAAACACCACTGAGCTGG 24
||||| ||| |||||
Db 25323 AAAAAAACCTCACTGAGCTGG 25300

Search completed: February 3, 2006, 21:26:00
Job time : 877 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 3, 2006, 10:53:33 ; Search time 203.111 Seconds
(without alignments)
820.326 Million cell updates/sec

Title: US-10-719-900-3

Perfect score: 25

Sequence: 1 aaaaacaaccaccactgagctggg 25

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 100%

Database : N Geneseq_21.*

1: geneseqn1980s.*

2: geneseqn1990s.*

3: geneseqn2000s.*

4: geneseqn2001as.*

5: geneseqn2001bs.*

6: geneseqn2002as.*

7: geneseqn2002bs.*

8: geneseqn2003as.*

9: geneseqn2003bs.*

10: geneseqn2003cs.*

11: geneseqn2003ds.*

12: geneseqn2004as.*

13: geneseqn2004bs.*

14: geneseqn2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	19.2	76.8	503	9	ACH44272 Human foe
C 2	18.8	75.2	42061	13	ABD33566 Human can
C 3	18.8	75.2	96898	10	ACH01384 Human BIV
C 4	18.6	74.4	90	8	ABZ09104 Human oil
C 5	18.6	74.4	90	10	ABZ78557 Tumour su
C 6	18.6	74.4	93	2	AAT24655 Human gen
C 7	18.6	74.4	382	10	ADD47465 Rat gene
C 8	18.6	74.4	540	4	ABA08904 Human oil
C 9	18.6	74.4	540	10	ADI21508 Novel hum
C 10	18.6	74.4	540	13	ADS98314 Protein E
C 11	18.6	74.4	591	4	AAH32100 Human oil
C 12	18.6	74.4	762	6	ABL39719 Human NS
C 13	18.6	74.4	777	6	ABZ42980 Human GPC
C 14	18.6	74.4	812	11	ACL1584 Rice abio
C 15	18.6	74.4	932	4	AAH32071 Human oil
C 16	18.6	74.4	950	8	ABX34731 Human mdd
C 17	18.6	74.4	951	10	ABZ77938 Human G p
C 18	18.6	74.4	1065	5	AAS15907 DNA encod
C 19	18.6	74.4	1069	6	ABK40204 Human G p

C 20	18.6	74.4	1232	13	ADS97926	Ad97926 Protein f
C 21	18.6	74.4	1329	6	ABN89112	Abn89112 Human GPC
C 22	18.6	74.4	1344	6	ABA95482	Aba95482 Human lam
C 23	18.6	74.4	1450	6	ABK92484	Abk92484 Human pro
C 24	18.6	74.4	2027	6	ABK95452	Abk95452 DNA encod
C 25	18.6	74.4	2028	6	ABK40203	Abk40203 Human G p
C 26	18.6	74.4	2312	8	ABZ588229	Abz588229 Human imm
C 27	18.6	74.4	2464	4	AAK68477	Aak68477 Human imm
C 28	18.6	74.4	2464	4	AAK68479	Aak68479 Human imm
C 29	18.6	74.4	2464	4	AAK04040	Aal04040 Human rep
C 30	18.6	74.4	2464	4	AAK04041	Aal04041 Human rep
C 31	18.6	74.4	2775	12	ADI16291	Adi16291 Human nuc
C 32	18.6	74.4	3009	12	ADO64688	Ado64688 Novel hum
C 33	18.6	74.4	3029	12	ADQ24417	Adq24417 Human sof
C 34	18.6	74.4	3030	12	ADI16310	Adi16310 Human nuc
C 35	18.6	74.4	3150	13	ADR25447	Adr25447 Breast ca
C 36	18.6	74.4	3190	3	AAZ57865	Aaz57865 Protein r
C 37	18.6	74.4	3226	14	AXO5939	Adx05939 Cyclin-de
C 38	18.6	74.4	3512	13	ACN42607	Acn42607 Human dta
C 39	18.6	74.4	4281	13	ADR07844	Adr07844 Full leng
C 40	18.6	74.4	9840	4	AAK67429	Aak67429 Human imm
C 41	18.6	74.4	16578	10	ADI22003	Adi22003 Novel hum
C 42	18.6	74.4	16578	13	ADS98483	Ads98483 Protein f
C 43	18.6	74.4	39119	8	ADA98641	Abz74034 Secreted
C 44	18.6	74.4	39119	8	ADA98641	Ada98641 Human sec
C 45	18.6	74.4	39119	10	ADC20764	Adc20764 Human sec
C 46	18.6	74.4	39119	10	ABZ67621	Abz67621 Human sec
C 47	18.6	74.4	53099	6	ABS98185	Abs98185 Human mul
C 48	18.6	74.4	86950	14	ADX98572	Adx98572 Human HTO
C 49	18.6	74.4	122923	11	ACN44026	Acn44026 Human gen
C 50	18.6	74.4	128993	8	ACF62749	Acf62749 Cancer ba
C 51	18.6	74.4	128993	8	ADB20868	Adb20868 MRP1 base
C 52	18.6	74.4	128993	10	ADB87957	Adb87957 Human UGT
C 53	18.6	74.4	128993	10	ADB96940	Adb96940 Human MDR
C 54	18.6	74.4	128993	10	ADB92131	Adb92131 Human MDR
C 55	18.6	74.4	177380	8	ACF62751	Acf62751 Cancer ba
C 56	18.6	74.4	177380	8	ADB20870	Adb20870 MRP1 base
C 57	18.6	74.4	177380	10	ADB87959	Adb87959 Human UGT
C 58	18.6	74.4	177380	10	ADB96942	Adb96942 Human MDR
C 59	18.6	74.4	177380	10	ADB92133	Adb92133 Human MDR
C 60	18.6	74.4	241748	14	ADZ13116	Adz13116 Murine ca
C 61	18.4	73.6	2097	8	ACA52966	ACA52966 Prokaryot
C 62	18.4	73.6	47756	12	ADQ59512	Adq59512 Human can
C 63	18.4	73.6	57036	14	ADZ13891	Adz13891 Human can
C 64	18.2	72.8	407	3	AAC22116	Aac22116 Human sec
C 65	18.2	72.8	598	4	ABA60937	Ab60937 Human foe
C 66	18.2	72.8	598	4	AAI40833	Aai40833 Probe #95
C 67	18.2	72.8	598	4	ABA28906	Ab28906 Probe #73
C 68	18.2	72.8	598	4	AAK35116	Aak35116 Human bon
C 69	18.2	72.8	598	4	AAK09227	Aak09227 Human bra
C 70	18.2	72.8	598	4	ABS34868	Abs34868 Human liv
C 71	18.2	72.8	598	6	ABS09572	Abs09572 Human gen
C 72	18.2	72.8	4624	4	ABL22501	Ab122501 Drosophil
C 73	18.2	72.8	110000	14	ADZ42274_2	Continuation (3 of
C 74	18.2	72.8	326014	6	ABK89296	Abk89296 Human gen
C 75	18.2	72.8	326014	12	ADQ94981	Adq94981 Human kin
C 76	17.8	71.2	108	4	ABA71547	Ab71547 Human foe
C 77	17.8	71.2	108	4	AAI51824	Aai51824 Probe #20
C 78	17.8	71.2	108	4	AAK45906	Aak45906 Human bon
C 79	17.8	71.2	108	4	AAK19876	Aak19876 Human bra
C 80	17.8	71.2	108	4	ABS45604	Abs45604 Human liv
C 81	17.8	71.2	108	6	ABS20191	Abs20191 Human gen
C 82	17.8	71.2	352	4	AAS26480	Aas26480 Human cdn
C 83	17.8	71.2	352	4	AAI62482	Aai62482 Human bre
C 84	17.8	71.2	352	4	AAK00448	Aal00448 Human rep
C 85	17.8	71.2	352	8	ABX73821	Abx73821 Human nov
C 86	17.8	71.2	360	4	ABA59020	Aba59020 Human foe
C 87	17.8	71.2	360	4	AAI38747	Aai38747 Probe #74
C 88	17.8	71.2	360	4	AAK32941	Aak32941 Human bon
C 89	17.8	71.2	360	4	AAK07190	Aak07190 Human bra
C 90	17.8	71.2	360	4	ABS32670	Abs32670 Human liv
C 91	17.8	71.2	360	6	ABS07747	Abs07747 Human gen
C 92	17.8	71.2	435	5	AAS67401	Aas67401 DNA encod


```
XX WO2004058146-A2.
PN
XX
XX 15-JUL-2004.
PD
XX
XX 15-DEC-2003; 2003WO-US040081.
PF
XX
XX 17-DEC-2002; 2002US-00322281.
PR
XX
XX (SAGR-) SAGRES DISCOVERY INC.
PA
XX
XX Morris DW, Malandro MS;
PI
XX
XX WPI; 2004-499109/47.
DR
XX
XX Novel human cancer associated protein encoded within open reading frame
PT
PT of cancer associated gene, useful as targets for diagnosing cancer.
XX
XX Claim 16; SEQ ID NO 770; 182pp; English.
PS
XX
XX The invention relates to cancer-associated proteins (CAP) and the cancer-
CC
CC associated (CA) nucleic acids encoding them. The invention also relates
CC
CC to a method for treating cancers involving administering to a patient an
CC
CC inhibitor of CAP, and a method of screening for anticancer activity in a
CC
CC potential drug involving providing a cell that expresses a CA gene,
CC
CC contacting a tissue sample derived from a cancer cell with an anticancer
CC
CC drug candidate and monitoring the effect of the anticancer drug candidate
CC
CC on expression of the CA gene. The CAP proteins are useful for detecting
CC
CC cancer associated with expression of a CAP protein in a test cell sample
CC
CC and for screening for a bioactive agent capable of modulating the
CC
CC activity of a CAP protein. The CA nucleic acids are useful for diagnosing
CC
CC cancer, involving determining the expression of a CA nucleic acid in a
CC
CC tissue. This sequence represents a human CA gene of the invention. Note:
CC
CC The sequence data for this patent did not form part of the printed
CC
CC specification, but was obtained in electronic format directly from WIPO
CC
CC at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 42061 BP; 12193 A; 8355 C; 8658 G; 12815 T; 0 U; 40 Other;
SQ
Query Match 75.2%; Score 18.8; DB 13; Length 42061;
Best Local Similarity 90.9%; Pred. No. 3.9e+02;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 AAAAAACACACCACCTGAGCT 22
Db 10151 AAAAAACACACCACTGAGCT 10130
RESULT 3
ACH01384
ID ACH01384 standard; DNA; 96898 BP.
XX
XX ACH01384;
AC
XX
XX 22-APR-2004 (first entry)
DT
XX
XX Human BIVM gene.
DE
XX
XX Human; mouse; zebrafish; chicken; sea urchin; BIVM; immunostimulant;
KW
KW basic immunoglobulin-like variable motif containing; cytostatic; cancer;
KW
KW gene; ds.
XX
XX Homo sapiens.
OS
XX
XX WO2003089595-A2.
PN
XX
XX 30-OCT-2003.
PD
XX
XX 16-APR-2003; 2003WO-US011768.
PF
XX
XX 16-APR-2002; 2002US-0373146P.
PR
XX
XX (UYSF-) UNIV SOUTH FLORIDA.
PA
```

```
XX Litman GW, Hawke NA, Yoder JA, Eason DD;
PI
XX WPI; 2003-854106/79.
DR
XX
XX New isolated or recombinant basic immunoglobulin-like variable motif-
PT
PT containing (BIVM) polynucleotides and polypeptides, useful for inducing
PT
PT an immune response or protective immunity against cancer, and for
PT
PT detecting Giardia.
XX
XX Claim 1; Page 68-123; Opp; English.
PS
XX
XX The present invention relates to isolated forms of basic, immunoglobulin-
CC
CC like variable motif-containing (BIVM) genes and encoded proteins. These
CC
CC can be used to induce an immune response or protective immunity against
CC
CC cancer. The present sequence is a coding sequence fragment shown in the
CC
CC exemplification of the invention
XX
XX Sequence 96898 BP; 27199 A; 18553 C; 20089 G; 31056 T; 0 U; 1 Other;
SQ
Query Match 75.2%; Score 18.8; DB 10; Length 96898;
Best Local Similarity 90.9%; Pred. No. 4.4e+02;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 AAAAAACACACCACCTGAGCT 22
Db 5945 AAAAAACACACCACCTGAGCT 5966
RESULT 4
ABZ09104/c
ID ABZ09104 standard; DNA; 90 BP.
XX
XX ABZ09104;
AC
XX
XX 16-JAN-2003 (first entry)
DT
XX
XX Human oligonucleotide SEQ ID 264.
DE
XX
XX Human; tumour suppressor; virucide; cytostatic; nootropic;
KW
KW neuroprotective; neuroleptic; gene therapy; tumour suppression;
KW
KW tumour reversion; apoptosis; viral resistance; viral infection;
KW
KW cell degeneration; Alzheimer's disease; schizophrenia; cancer; ds.
XX
XX Homo sapiens.
OS
XX
XX FR2822475-A1.
PN
XX
XX 27-SEP-2002.
PD
XX
XX 20-MAR-2002; 2002FR-00003459.
PF
XX
XX 13-FEB-2001; 2001FR-00001925.
PR
XX
XX (MOLE-) MOLECULAR ENGINES LAB SA.
PA
XX
XX Telerman A, Anson R, Tuijnder M, Susini L;
PI
XX
XX WPI; 2003-032204/03.
DR
XX
XX New human nucleic acid, useful for diagnosis, prognosis and treatment,
PT
PT e.g. of tumors, also related vectors, transformed cell, polypeptides and
PT
PT antibodies.
XX
XX Disclosure; Page 78; 189pp; French.
PS
XX
XX The present invention relates to human oligonucleotides (ABZ08941-
CC
CC ABZ09860). The expression of the oligonucleotides is implicated in tumour
CC
CC suppression or reversion, apoptosis and/or viral resistance. The
CC
CC oligonucleotides are useful for preventing and/or treating viral
CC
CC infection, tumour development and cell degeneration (e.g. Alzheimer's
CC
CC disease and schizophrenia), especially cancer
XX
```

```

SQ Sequence 90 BP; 30 A; 16 C; 13 G; 31 T; 0 U; 0 Other;
  Query Match          74.4%; Score 18.6; DB 8; Length 90;
  Best Local Similarity 84.0%; Pred. No. 1.8e+02;
  Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
  QY 1 AAAAAACAACCACTGAGCTGGG 25
     |||||
  Db 43 AAAAAACAACCACTGAAATGTG 19

RESULT 5
ABZ78557/c
ID ABZ78557 standard; DNA; 90 BP.
XX
AC ABZ78557;
XX
DT 24-APR-2003 (first entry)
XX
XX Tumour suppression-related sequence, SEQ ID 264.
XX
XX Cytostatic; virucide; apoptotic; gene therapy; tumour suppression;
KW tumour reversion; apoptosis; virus resistance; viral infection; tumour;
KW cell degenerative disease; ds.
XX
OS Unidentified.
XX
XX WO200264731-A2.
XX
XX 22-AUG-2002.
XX
XX 13-FEB-2002; 2002WO-FR000543.
XX
XX 13-FEB-2001; 2001FR-00001925.
XX
XX (MOLE-) MOLECULAR ENGINES LAB.
XX
XX Telerman A, Amson R, Tuijnder M, Susini L;
XX WPI; 2003-058286/05.
XX
XX New nucleic acid encoding a translationally controlled tumor protein,
PT useful for treating, preventing and diagnosing viral, tumor or
PT degenerative diseases.
XX
XX Disclosure; Page; 45pp; French.
XX
XX The present invention relates to novel nucleic acid sequences (ABZ78294-
CC ABZ79313), which are involved in the molecular pathways of tumour
CC suppression, tumour reversion, apoptosis and/or virus resistance. The
CC sequences are also useful for treatment or prevention of viral, tumour
CC and cell degenerative diseases, and also for diagnosis and prognosis of
CC these diseases. Note: The sequence data for this patent is not
CC represented in the printed specification but is based on sequence
CC information supplied by the European Patent Office
XX
SQ Sequence 90 BP; 30 A; 16 C; 13 G; 31 T; 0 U; 0 Other;
  Query Match          74.4%; Score 18.6; DB 10; Length 90;
  Best Local Similarity 84.0%; Pred. No. 1.8e+02;
  Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
  QY 1 AAAAAACAACCACTGAGCTGGG 25
     |||||
  Db 43 AAAAAACAACCACTGAAATGTG 19

RESULT 6
AAT24655
ID AAT24655 standard; cDNA to mRNA; 93 BP.
XX
AC AAT24655;
XX

SQ Sequence 93 BP; 34 A; 13 C; 16 G; 30 T; 0 U; 0 Other;
  Query Match          74.4%; Score 18.6; DB 2; Length 93;
  Best Local Similarity 84.0%; Pred. No. 1.8e+02;
  Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
  QY 1 AAAAAACAACCACTGAGCTGGG 25
     |||||
  Db 48 AAAAAACAACCACTGAAATGTG 72

RESULT 7
ADD47465
ID ADD47465 standard; DNA; 382 BP.
XX
AC ADD47465;
XX
DT 02-DEC-2004 (revised)
DT 29-JAN-2004 (first entry)
XX
XX Rat gene AA852046, SEQ ID NO 13160.
XX
XX Rat; ds; gene; pain; neuronal tissue; gene therapy;
KW spinal segmental nerve injury; chronic constriction injury; CCI;
KW spared nerve injury; SNI; Chung.
XX
XX Rattus norvegicus.
```


CC autoimmune disease or accidental damage. The polypeptides and nucleotides
CC may also be used in the diagnosis of the above conditions, and in drug
CC screening techniques. The present sequence represents a cDNA encoding a
CC novel human polypeptide of the invention

XX
SQ Sequence 540 BP; 139 A; 132 C; 160 G; 109 T; 0 U; 0 Other;

Query Match 74.4%; Score 18.6; DB 4; Length 540;

Best Local Similarity 84.0%; Pred. No. 2.4e+02;

Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAACACCACTGAGCTGGG 25

|||||

Db 323 AAAAAACCACTGCGGTG 347

RESULT 9

AD121508

ID AD121508 standard; cDNA; 540 BP.

XX

AC AD121508;

XX

DT 15-APR-2004 (first entry)

XX

DE Novel human expressed sequence tag, EST #207.

XX

KW forensic; nutritional source; damaged tissue; diseased tissue;

myeloid cell disorder; lymphoid cell disorder;

bone cartilage tissue growth; tendon tissue growth;

ligament tissue growth; nerve tissue growth; regeneration; wound healing;

tissue repair; tissue replacement; burn; incision; ulcer; cancer; human;

SS; Gene; EST; expressed sequence tag.

XX

OS Homo sapiens.

XX

PN WO2003025148-A2.

XX

PD 27-MAR-2003.

XX

PF 19-SEP-2002; 2002WO-US029964.

XX

PR 19-SEP-2001; 2001US-0323739P.

PR

PR 13-SEP-2002; 2002US-00323739.

XX

PA (HYSE-) HYSEQ INC.

XX

PI Tang YT, Asundi V, Goodrich RW, Ren F, Zhang J, Zhao QA, Wang J;

Ghosh M, Xue AJ, Wehrman T, Weng G, Zhou P, Drmanac RT, Wang D;

PI Haley-Vicente D;

XX

DR WPI; 2003-354603/33.

DR

P-PSDB; AD121728.

XX

PT New polynucleotides and secreted proteins, useful for treating myeloid or
PT lymphoid cell disorders, in bone cartilage, tendon, ligament and nerve
PT tissue growth or regeneration, in wound healing, and in tissue repair and
PT replacement.

XX

PS Example 2; SEQ ID NO 759; 156pp; English.

XX

CC The invention relates to an isolated polynucleotide encoding a

polypeptide with biological activity. The polynucleotides and

polypeptides are useful in diagnostics, forensics, gene mapping,

CC identification of mutations responsible for genetic disorders and other

traits, to assess biodiversity, as nutritional sources or supplements.

CC The polynucleotides may also be used as molecular weight markers,

CC chromosome markers or map related gene positions, or as an antigen to

raise anti-DNA antibodies or elicit immune response. The polypeptides are

CC useful for raising antibodies, as markers for tissues in which the

corresponding polypeptide is expressed, for re-engineering damaged or

CC diseased tissues, for treating myeloid or lymphoid cell disorders, in

CC bone cartilage, tendon, ligament and/or nerve tissue growth or

regeneration, in wound healing, in tissue repair and replacement, in

CC

CC healing of burns, incisions and ulcers, and in treating cancer. The
CC present sequence represents a novel human expressed sequence tag, EST.

XX

SQ Sequence 540 BP; 139 A; 132 C; 160 G; 109 T; 0 U; 0 Other;

Query Match 74.4%; Score 18.6; DB 10; Length 540;

Best Local Similarity 84.0%; Pred. No. 2.4e+02;

Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAACACCACTGAGCTGGG 25

|||||

Db 323 AAAAAACCACTGCGGTG 347

RESULT 10

ADS98314

ID ADS98314 standard; DNA; 540 BP.

XX

AC ADS98314;

XX

DT 30-DEC-2004 (first entry)

XX

DE Protein factor discovery related isolated human contig DNA, SEQ ID 578.

XX

KW antiinflammatory; cytostatic; antimicrobial; gene therapy; inflammation;

leukaemia; nervous system disorder; infection; gene; ds.

XX

OS Homo sapiens.

XX

PN WO2004087874-A2.

XX

PD 14-OCT-2004.

XX

PF 24-MAR-2004; 2004WO-US009202.

XX

PR 28-MAR-2003; 2003US-0458824P.

XX

PA (NUVE-) NUVELO INC.

PA (DRNA/) DRMANAC R T.

XX

PI Tang YT, Zhou P, Wang J, Wang ZW, Hu T;

XX

WPI; 2004-737686/72.

DR

P-PSDB; ADS98654.

XX

PT New polynucleotides encoding a polypeptide with biological activity,
PT useful for treating inflammation, leukemias, nervous system disorders, or
PT infections.

XX

PS Example 2; SEQ ID NO 578; 253pp; English.

XX

CC The invention relates to a novel isolated polynucleotide comprising any
of the 235 nucleotide sequences described in the specification. The

CC invention further comprises: an isolated polynucleotide encoding a

polypeptide with biological activity, where the polynucleotide hybridizes

CC to one of the 235 novel polynucleotides under stringent hybridization

conditions, or having greater than about 99% sequence identity with the

CC novel polynucleotide; a vector comprising a novel polynucleotide; an

expression vector comprising the novel polynucleotide; a host cell

CC genetically engineered to comprise the novel polynucleotide, which can be

operatively associated with a regulatory sequence that modulates

CC expression of the polynucleotide in the host cell; an isolated

polypeptide encoded by the novel polynucleotide, or a polynucleotide

CC hybridizing under stringent conditions to the novel polynucleotide; a

composition comprising the polypeptide and a carrier; an antibody

CC directed against the polypeptide; a method for detecting the novel

polynucleotide in a sample; a method for detecting the polypeptide in a

CC sample; a method for identifying a compound that binds to the polypeptide

CC; a method for producing the polypeptide; an isolated polypeptide

CC comprising any of the 235 amino acid sequences described in the

CC specification; and a collection of polynucleotides comprising of at least

one of the polynucleotides cited above. The polypeptides and

CC polynucleotides of the invention have antiinflammatory, cytostatic, and

CC disease, epilepsy, angina, neurodegeneration, diabetes, anxiety,
CC depression, schizophrenia, viral disease, gastric ulcers, stroke,
CC Alzheimer's disease and as a contraceptive
XX
SQ Sequence 762 BP; 244 A; 138 C; 187 G; 178 T; 0 U; 15 Other;
Query Match 74.4%; Score 18.6; DB 6; Length 762;
Best Local Similarity 84.0%; Pred. No. 2.6e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAAACACCACTGAGCTGGG 25
|||||
Db 711 AAAAAAACACCACTGAGCTGG 735
|||||

RESULT 13
ID ABZ42980/c
XX ABZ42980 standard; DNA; 777 BP.
XX
AC ABZ42980;
XX
DT 06-MAR-2003 (first entry)
XX
DE Human GPCR polynucleotide SEQ ID NO 221.
XX
KW Human; GPCR; G protein coupled receptor; signal transduction; olfactory;
KW drug development; gustatory; taste; fragrance; gene; ds.
XX
OS Homo sapiens.
XX
PN WO200216548-A2.
XX
PD 28-FEB-2002.
XX
PF 30-JUL-2001; 2001WO-IB001446.
XX
PR 04-AUG-2000; 2000JP-00237818.
PR 13-FEB-2001; 2001JP-00034433.
XX
PA (NISC-) JAPAN SCI & TECHNOLOGY CORP.
XX
PI Haga T, Takeda S, Mitaku S;
XX
XX WPI; 2002-304118/34.
DR P-PSDB; ABP95706.
XX
PT Database global search for G protein-coupled receptors, proteins and
PT encoded genes for studying in vivo signal transduction mechanism and
PT identifying targets for drug development.
XX
PS Claim 9; SEQ ID NO 221; 97pp + Sequence Listing; Japanese.
XX
CC The invention relates to a method for screening G protein-coupled
CC receptor (GPCR) genes (ABZ42870-ABZ43216) and/or GPCR proteins (ABP95596-
CC ABP95942) by extracting open-reading frames containing 6-8 transmembrane
CC domains with 250-1000 amino acid residues to give a gene homologous with
CC a known GPCR gene. The receptor proteins and encoded genes are useful for
CC studying in vivo signal transduction mechanism and identifying targets
CC for drug development e.g. based on olfactory and gustatory receptors in
CC form of agonists and antagonists by screening intrinsic and extrinsic
CC ligands as bitter taste inhibitors, taste enhancers and fragrance
CC improvers. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 777 BP; 145 A; 241 C; 194 G; 197 T; 0 U; 0 Other;
Query Match 74.4%; Score 18.6; DB 6; Length 777;
Best Local Similarity 84.0%; Pred. No. 2.6e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAAACACCACTGAGCTGGG 25
|||||

Db 544 AAAAAAACCACTGTGGGTG 520
RESULT 14
ID ACL31584/c
XX ACL31584 standard; cDNA; 812 BP.
XX
AC ACL31584;
XX
DT 02-JUN-2005 (first entry)
XX
DE Rice abiotic stress response related polynucleotide SEQ ID NO:10147.
XX
KW ss; abiotic stress tolerance; transgenic plant; plant; cereal;
KW agriculture.
XX
OS Oryza sativa.
XX
PN WO2003008540-A2.
XX
PD 30-JAN-2003.
XX
PF 21-JUN-2002; 2002WO-US019668.
XX
PR 22-JUN-2001; 2001US-0300112P.
PR 24-AUG-2001; 2001US-0314662P.
PR 26-SEP-2001; 2001US-0325277P.
PR 21-NOV-2001; 2001US-0332132P.
XX
PA (SYGN) SYNGENTA PARTICIPATIONS AG.
XX
PI Kreps J, Briggs SP, Cooper B, Glazebrook J, Goff SA, Katagiri P;
PI Moughamer T, Provart N, Ricke D, Zhu T;
XX
DR WPI; 2003-248011/24.
XX
PT New stress-responsive nucleic acid, useful for altering the
PT responsiveness of a plant, e.g. cereal, to an abiotic stress such as cold
PT stress, salt stress or osmotic stress.
XX
PS Disclosure; SEQ ID NO 10147; 89pp; English.
XX
CC The invention relates to novel abiotic stress responsive polynucleotides
CC and polypeptides. Also disclosed are vectors, expression cassettes, host
CC cells, and plants containing such polynucleotides. Also disclosed are
CC methods for using the polynucleotides and polypeptides to alter the
CC responsiveness of a plant to abiotic stress. The invention is useful in
CC agriculture. The nucleic acid is useful for determining whether a test
CC plant has been exposed to an abiotic stress condition. It is also useful
CC for selecting an agent that alters abiotic stress regulated
CC polynucleotide expression in a plant cell, and to identify a homolog or
CC ortholog to an abiotic stress responsive polynucleotide. The nucleic acid
CC molecule and the polypeptide encoded by it are useful in altering the
CC responsiveness of a plant to an abiotic stress, such as cold stress, salt
CC stress, osmotic stress or any of their combinations. The present sequence
CC is used in the exemplification of the invention
XX
SQ Sequence 812 BP; 171 A; 212 C; 225 G; 204 T; 0 U; 0 Other;
Query Match 74.4%; Score 18.6; DB 11; Length 812;
Best Local Similarity 84.0%; Pred. No. 2.6e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAAACCACTGAGCTGGG 25
|||||
Db 509 AAGAATCACTTACTGAGCTGGG 485
|||||

RESULT 15
ID AAH32071/c
XX AAH32071 standard; DNA; 932 BP.
XX
AC AAH32071;

XX 30-JUL-2001 (first entry)
XX Human olfactory receptor polynucleotide, SEQ ID NO: 644.
DE
XX
XX Human; olfactory receptor; OR; primary scent determination;
KW secondary scent determination; polypeptide library; odour receptor;
KW scent profile; scent fingerprint; scent representation; ds.
XX
XX Homo sapiens.
OS
XX WO200127158-A2.
XX
XX 19-APR-2001.
XX
XX 06-OCT-2000; 2000WO-US027582.
XX
XX 08-OCT-1999; 99US-0158615P.
XX
XX 24-FEB-2000; 2000US-0184809P.
XX
XX (DIGI-) DIGISCENTS.
XX (YEDA) YEDA RES & DEV CO LTD.
PA
XX Bellenson J, Smith D, Lancet D, Glusman G, Fuchs T, Yanai I;
XX
XX WPI; 2001-290713/30.
XX
XX New polynucleotides which encode polypeptides involved in olfactory
PT sensation for identifying olfactory agonists and antagonists.
PT
XX Claim 8; Page 439; 1857pp; English.
XX
XX The present sequence is one of a number of isolated polynucleotides which
CC encode polypeptides involved in olfactory sensation. The polynucleotides
CC can be used in screening for olfactory agonists and antagonists. The
CC methods allow for the determination of primary scents and the
CC identification of the odour receptors used to detect these primary
CC scents. The methods also enable determination of secondary scents and the
CC identification of combinations of odour receptors that are involved in
CC detecting such secondary scents. This enables the construction of a scent
CC representation (also called a scent fingerprint or scent profile), which
CC may be used to re-create and edit scents. Libraries of olfactory
CC receptors are useful for determining the interaction pattern of a
CC composition with the receptors, and can be used for determining
CC differences in the olfactory faculties of different individuals
XX
XX Sequence 932 BP; 179 A; 284 C; 236 G; 233 T; 0 U; 0 Other;
SQ
Query Match 74.4%; Score 18.6; DB 4; Length 932;
Best Local Similarity 84.0%; Pred. No. 2.6e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 AAAAAACACACCACCACTGAGCTGGG 25
DB 744 AAAAAACACACCACCACTGAGCTGGG 720
RESULT 16
ABX34731/c
ID ABX34731 standard; cDNA; 950 BP.
XX
XX AC ABX34731;
XX
XX 13-FEB-2003 (first entry)
XX
XX Human mddt cDNA SEQ ID 292.
DE
XX
XX MDDT; human; disease detection and treatment molecule polypeptide;
KW anti-inflammatory; immunosuppressive; osteopathic; cycostatic; anti-HIV;
KW haemostatic; nephrotropic; antianaemic; antipsoriatic; hepatotropic;
KW gene therapy; protein replacement therapy; cell proliferative disorder;
KW cancer; adenocarcinoma; leukaemia; lymphoma; melanoma; myeloma; sarcoma;
KW anaemia; Crohn's disease; acquired immunodeficiency syndrome; AIDS;

KW Goodpasture's syndrome; inflammation; osteoporosis; thrombocytopaenia;
KW psoriasis; hepatitis; gene; ss.
XX
XX Homo sapiens.
OS
XX WO200279449-A2.
XX
XX 10-OCT-2002.
XX
XX 27-MAR-2002; 2002WO-US009944.
XX
XX 28-MAR-2001; 2001US-0279619P.
XX
XX 29-MAR-2001; 2001US-0280067P.
XX
XX 29-MAR-2001; 2001US-0280068P.
XX
XX 16-MAY-2001; 2001US-0291280P.
XX
XX 17-MAY-2001; 2001US-0291829P.
XX
XX 17-MAY-2001; 2001US-0291849P.
XX
XX 19-JUN-2001; 2001US-0299428P.
XX
XX 20-JUN-2001; 2001US-0299776P.
XX
XX 20-JUN-2001; 2001US-0300001P.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Daffo A, Jones AL, Tran AB, Dahl CR, Gietzen D, Chinn J;
XX Dufour GE, Hillman JL, Yu JY, Tuason O, Yap PE, Amshay SR;
XX Daugherty SC, Dam TC, Liu TF, Nguyen DA, Kleefeld Y, Gerstin EH;
XX Paralta CH, David MH, Lewis SA, Chen AJ, Panzer SR, Harris B;
XX Flores V, Marwaha R, Lo A, Lan RV, Urashka ME;
XX WPI; 2003-058431/05.
XX P-PSDB; ABU11741.
XX
XX New purified disease detection and treatment molecule proteins and
PT polynucleotides, useful for diagnosing, treating or preventing cancers
PT (e.g. leukemia or sarcoma), anaemia, Crohn's disease, AIDS, osteoporosis
PT or hepatitis.
XX
XX Claim 1; SEQ ID NO 292; 339pp + Sequence Listing; English.
XX
XX This invention describes a novel disease detection and treatment molecule
CC polypeptide (MDDT) which has anti-inflammatory, immunosuppressive,
CC osteopathic, cycostatic, anti-HIV, haemostatic, nephrotropic,
CC antianaemic, antipsoriatic and hepatotropic activity. The polynucleotides
CC and the polypeptides of the invention can be used for gene therapy,
CC protein replacement therapy and are useful for treating a variety of
CC diseases or conditions. These polypeptides or polynucleotides are
CC particularly useful for diagnosing, treating or preventing cell
CC proliferative disorders (e.g. cancers including adenocarcinoma,
CC leukaemia, lymphoma, melanoma, myeloma or sarcoma), anaemia, Crohn's
CC disease, acquired immunodeficiency syndrome (AIDS), Goodpasture's
CC syndromes, inflammation, osteoporosis, thrombocytopaenia, psoriasis or
CC hepatitis. ABX34440-ABX34835 encode the MDDT polypeptides represented in
CC ABU1450-ABU11845, described in the disclosure of the invention. NOTE:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 950 BP; 173 A; 300 C; 239 G; 238 T; 0 U; 0 Other;
SQ
Query Match 74.4%; Score 18.6; DB 8; Length 950;
Best Local Similarity 84.0%; Pred. No. 2.6e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 AAAAAACACACCACCACTGAGCTGGG 25
DB 690 AAAAAACACACCACCACTGAGCTGGG 666
RESULT 17
ABX77938/c
ID ABX77938 standard; DNA; 951 BP.
XX
XX AC ABX77938;

XX 17-APR-2003 (first entry)
 DT Human G protein coupled receptor coding sequence SEQ ID 133.
 XX
 DE Human; gene; anorectic; antiasthmatic; antidiabetic; hypotensive;
 XX antiparkinsonian; nootropic; neuroprotective; tranquiliser;
 KW antirheumatic; antiinflammatory; osteopathic; cardiant; neuroleptic;
 KW antithratic; gene therapy; olfactory G protein-coupled receptor; GPCR;
 KW infection; obesity; diabetes; hypertension; malnutrition;
 KW Parkinson's disease; Alzheimer's disease; Korsakoff's psychosis; anxiety;
 KW rheumatoid arthritis; chronic obstructive pulmonary disease;
 KW osteoporosis; asthma; myocardial infarction; schizophrenia;
 KW osteoarthritis; ds.
 XX
 XX Homo sapiens.
 OS
 XX WO2003000735-A2.
 PN
 XX 03-JAN-2003.
 PD
 XX 24-JUN-2002; 2002WO-IB002481.
 XX
 XX 26-JUN-2001; 2001US-0301095P.
 PR
 XX 06-NOV-2001; 2001US-0332758P.
 XX
 XX (DECO-) DECODE GENETICS EHF.
 PA
 XX Martinez RAM, Sigurdsson GT;
 XX
 XX WPI; 2003-175284/17.
 DR
 XX P-PSDB; ABR01637.
 XX
 XX New olfactory G protein-coupled receptor gene nucleic acid and
 PT polypeptide, useful for diagnosing or treating a disease or condition
 PT associated with GPCR, e.g. Obesity, diabetes, hypertension, malnutrition
 PT or Alzheimer's disease.
 XX
 XX Claim 1; Page 87; 383pp; English.
 PS
 XX The present invention relates to novel human olfactory G protein-coupled
 CC receptors (GPCR) and their coding sequences (AB277872-AB277986 and
 CC ABR01571-ABR01685). The GPCRs and coding sequences are useful for
 CC diagnosing or treating a disease or condition associated with GPCR, e.g.
 CC infections, obesity, diabetes, hypertension, malnutrition, Parkinson's
 CC disease, Alzheimer's disease, Korsakoff's psychosis, anxiety, rheumatoid
 CC arthritis, chronic obstructive pulmonary disease, osteoporosis, asthma,
 CC myocardial infarction, schizophrenia, or osteoarthritis
 XX
 XX Sequence 951 BP; 177 A; 298 C; 235 G; 241 T; 0 U; 0 Other;
 SQ
 Query Match 74.4%; Score 18.6; DB 10; Length 951;
 Best Local Similarity 84.0%; Pred. No. 2.6e+02;
 Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 1 AAAAAACACCACCTGAGCTGG 25
 Db 766 AAAAAACACCACCTGTCGGTG 742
 RESULT 18
 AAS15907/c
 ID AAS15907 standard; cDNA; 1065 BP.
 XX
 AC AAS15907;
 XX
 DT 25-JAN-2002 (first entry)
 XX
 XX DNA encoding G-protein coupled receptor (GCREC) #11.
 DE
 XX G-protein coupled receptor; GCREC; vaccine; gene therapy;
 KW cell proliferation disorder; cancer; arteriosclerosis;
 KW neurological disorder; epilepsy; stroke; cardiovascular disorder;
 KW

KW hypertension; ischaemic heart disease; gastrointestinal disorder;
 KW anorexia; peptic ulcer; autoimmune disorder; inflammatory disorder;
 KW diabetes; osteoporosis; psoriasis; metabolic disorder; obesity;
 KW schizophrenia disorder; neuroskeletal disorder; ss.
 XX
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FT CDS 1..1065
 FT /*tag= a
 FT /product= "GCREC 11"
 FT /note= "G-protein coupled receptor 11"
 XX
 PN WO200166742-A2.
 XX
 XX 13-SEP-2001.
 PD
 XX 01-MAR-2001; 2001WO-US006814.
 PF
 XX 03-MAR-2000; 2000US-0186854P.
 PR
 XX 10-MAR-2000; 2000US-0188384P.
 PR
 XX 17-MAR-2000; 2000US-0190453P.
 PR
 XX 20-MAR-2000; 2000US-0190730P.
 XX
 XX (INCY-) INCYTE GENOMICS INC.
 PA
 XX Lal P, Tang YT, Patterson C, Yao MG, Shih LL, Tribouley CM;
 PI Lu DAM, Yue H, Khan FA, Policky JL, Au-Young J, Yang J, Harland L;
 PI Walsh RT, Lo TP, Borowsky ML;
 XX
 XX WPI; 2001-656776/75.
 DR
 XX P-PSDB; AAU10310.
 XX
 XX Novel G-protein coupled receptor polypeptides, for treating and
 PT preventing autoimmune/inflammatory disorders, neurological disorders,
 PT cell proliferative disorders, cardiovascular disorders and viral
 PT infections.
 XX
 XX Claim 5; Page 136-137; 141pp; English.

XX The invention describes a novel isolated polypeptide, selected from a
 CC group of 21 G-protein coupled receptor polypeptides (GCREC) and useful in
 CC vaccines and gene therapy. The polypeptide (I) is useful for screening
 CC for agonist or antagonist of (I), compounds specifically binding to (I),
 CC or compounds that modulate the activity of (I). The polynucleotide
 CC encoding (I) is useful for screening a compound for effectiveness in
 CC altering expression of a target polynucleotide comprising (II), by
 CC exposing a sample comprising the target polynucleotide to a compound,
 CC detecting altered expression of the target polynucleotide, and comparing
 CC the expression of the target polynucleotide in the presence of varying
 CC amounts of compound and in the absence of the compound. (I) and (II) are
 CC useful for diagnosis, treatment and prevention of cell proliferative
 CC disorders (e.g. cancers, arteriosclerosis, atherosclerosis), neurological
 CC disorders (e.g. epilepsy, stroke, schizophrenic disorders and
 CC neuroskeletal disorders), cardiovascular disorders (e.g. anorexia,
 CC ischaemic heart disease), gastrointestinal disorders (e.g. diabetes mellitus,
 CC peptic ulcer), autoimmune/inflammatory disorders (e.g. diabetes mellitus,
 CC osteoporosis, psoriasis), and metabolic disorders such as obesity.
 CC Furthermore, the polynucleotide is useful; as primers for detecting
 CC single nucleotide polymorphisms; as elements in microarray, to monitor or
 CC measure protein-protein interactions, drug-target interactions, and gene
 CC expression profiles; to generate a transcript image of a tissue or cell
 CC type, and to generate hybridisation probes useful in mapping the
 CC naturally occurring genomic sequence. This sequence encodes G-protein
 CC coupled receptor 11, one of 21 GCREC proteins described in the method of
 CC the invention
 XX
 XX Sequence 1065 BP; 214 A; 327 C; 271 G; 253 T; 0 U; 0 Other;
 SQ

Query Match 74.4%; Score 18.6; DB 5; Length 1065;
 Best Local Similarity 84.0%; Pred. No. 2.7e+02;
 Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAACACACCACTGAGCTGGG 25
 DB 832 AAAAAACACCACTGAGCTGGG 808

RESULT 19
 ABK40204/c
 ID ABK40204 standard; cDNA; 1069 BP.
 XX AC ABK40204;
 XX 21-MAY-2002 (first entry)
 XX DE Human G protein-coupled receptor (GPCR) 9 cDNA #2.
 XX Human; G protein-coupled receptor; GPCR; cell signalling processing; ss;
 KW cardiomyopathy; atherosclerosis; developmental disease; immune disease;
 KW non insulin-dependent diabetes mellitus; NIDDM1; bacterial infection;
 KW fungal infection; protozoal infection; viral infection; adenocarcinoma;
 KW human immunodeficiency virus; HIV; cancer; lymphoma; anorexia; asthma;
 KW neurodegenerative disorder; Parkinson's disease; acute heart failure;
 KW hypotension; hypertension; osteoporosis; Crohn's disease; allergy; ulcer;
 KW multiple sclerosis; osteodystrophy; angina pectoris; schizophrenia; gene;
 KW myocardial infarction; neurological disorder; anxiety; manic depression;
 KW delirium; dementia; Huntington's disease; Alzheimer's disease; obesity;
 KW Tourette's syndrome; metabolic disorder.
 XX OS Homo sapiens.
 XX WO200202637-A2.
 XX 10-JAN-2002.
 XX 02-JUL-2001; 2001WO-US021174.
 XX 30-JUN-2000; 2000US-0215483P.
 XX 30-JUN-2000; 2000US-0215495P.
 XX 08-JUL-2000; 2000US-0216117P.
 XX 25-JUL-2000; 2000US-0220589P.
 XX 02-NOV-2000; 2000US-0245294P.
 XX 10-JAN-2001; 2001US-0260729P.
 XX 10-JAN-2001; 2001US-0260851P.
 XX 26-FEB-2001; 2001US-0271673P.
 XX (CURA-) CURAGEN CORP.
 XX Spaderna BK, Padigar M, Spytek KA, Casman S, Rastelli L;
 PI Mayankar U, Tchernev V;
 XX WPI; 2002-164524/21.
 XX P-PSDB; AAU85873.
 XX Novel polypeptides and nucleic acids of G protein coupled receptor useful
 PT for diagnosing and treating diseases e.g. anorexia, asthma, Parkinson's
 PT disease, acute heart failure and infections.
 XX Claim 9; Page 58; 202pp; English.
 XX The invention relates to human G protein coupled receptor (GPCR)
 CC polypeptides and the polynucleotides encoding them. GPCR polypeptides are
 CC useful for treating or preventing a GPCR-associated disorder which is
 CC related to cell signalling processing and metabolic pathway modulation,
 CC which includes cardiomyopathy and atherosclerosis in a human. The
 CC sequences are useful in treatment of patients suffering from
 CC developmental diseases, immune diseases, non insulin-dependent diabetes
 CC mellitus (NIDDM1), bacterial, fungal, protozoal and viral infections
 CC (e.g. infections caused by HIV), cancer (e.g. adenocarcinoma, lymphoma),
 CC anorexia, asthma, neurodegenerative disorders (e.g. Parkinson's disease),
 CC acute heart failure, hypertension, osteoporosis, Crohn's
 CC disease, multiple sclerosis, osteodystrophy, angina pectoris, myocardial
 CC infarction, ulcers, allergies, neurological disorders including anxiety,
 CC schizophrenia, manic depression, delirium, dementia, Huntington's
 CC disease, Alzheimer's disease, Tourette's syndrome, metabolic disorders

CC and obesity. This sequence represents cDNA encoding a GPCR polypeptide of
 CC the invention
 XX Sequence 1069 BP; 210 A; 315 C; 274 G; 270 T; 0 U; 0 Other;
 SQ Query Match 74.4%; Score 18.6; DB 6; Length 1069;
 Best Local Similarity 84.0%; Pred. No. 2.7e+02;
 Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 1 AAAAAACACCACTGAGCTGGG 25
 DB 778 AAAAAACCACTGAGCTGGG 754

RESULT 20
 ADS97926/c
 ID ADS97926 standard; DNA; 1232 BP.
 XX ADS97926;
 AC ADS97926;
 XX 30-DEC-2004 (first entry)
 DT Protein factor discovery related isolated human DNA, SEQ ID No 190.
 XX antiinflammatory; cytostatic; antimicrobial; gene therapy; inflammation;
 KW antileukemia; nervous system disorder; infection; gene; ds.
 KW Homo sapiens.
 XX WO2004087874-A2.
 XX 14-OCT-2004.
 XX 24-MAR-2004; 2004WO-US009202.
 XX 28-MAR-2003; 2003US-0458824P.
 XX (NUVE-) NUVELO INC.
 XX (DRMA/) DRMANAC R T.
 XX Tang YT, Zhou P, Wang J, Wang ZW, Hu T;
 XX WPI; 2004-737686/72.
 XX P-PSDB; ADS98161.
 XX New polynucleotides encoding a polypeptide with biological activity,
 PT useful for treating inflammation, leukemias, nervous system disorders, or
 PT infections.
 XX Claim 1; SEQ ID NO 190; 253pp; English.
 XX The invention relates to a novel isolated polynucleotide comprising any
 CC of the 235 nucleotide sequences described in the specification. The
 CC invention further comprises: an isolated polynucleotide encoding a
 CC polypeptide with biological activity, where the polynucleotide hybridizes
 CC to one of the 235 novel polynucleotides under stringent hybridization
 CC conditions, or having greater than about 99% sequence identity with the
 CC novel polynucleotide; a vector comprising a novel polynucleotide; an
 CC expression vector comprising the novel polynucleotide; a host cell
 CC genetically engineered to comprise the novel polynucleotide, which can be
 CC operatively associated with a regulatory sequence that modulates
 CC expression of the polynucleotide in the host cell; an isolated
 CC polypeptide encoded by the novel polynucleotide, or a polynucleotide
 CC hybridizing under stringent conditions to the novel polynucleotide; a
 CC composition comprising the polypeptide and a carrier; an antibody
 CC directed against the polypeptide; a method for detecting the novel
 CC polynucleotide in a sample; a method for detecting the polypeptide in a
 CC sample; a method for identifying a compound that binds to the polypeptide
 CC ; a method for producing the polypeptide; an isolated polypeptide
 CC comprising any of the 235 amino acid sequences described in the
 CC specification; and a collection of polynucleotides comprising of at least
 CC one of the polynucleotides cited above. The polypeptides and
 CC polynucleotides of the invention have antiinflammatory, cytostatic, and

CC antimicrobial activities. The novel polynucleotide may be used to treat
 CC disorders by gene therapy. The polypeptides and polynucleotides are
 CC useful for treating inflammation, leukaemias, nervous system disorders,
 CC or infections. This sequence represents one of the 235 novel isolated
 CC polynucleotides of the invention.

XX
 SQ Sequence 1232 BP; 252 A; 358 C; 315 G; 307 T; 0 U; 0 Other;
 Query Match 74.4%; Score 18.6; DB 13; Length 1232;
 Best Local Similarity 84.0%; Pred. No. 2.7e+02;
 Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 AAAAAACACCACTGAGCTGGG 25

Db 916 AAAAAACCACTGAGCTGGG 892

RESULT 21

ABN89112/C

ID ABN89112 standard; cDNA; 1329 BP.

XX AC ABN89112;

XX DT 28-AUG-2002 (first entry)

XX DE Human GPCR1 nucleotide sequence SEQ ID NO:1.

XX KW Human; GPCR; GPCR; G protein-coupled receptor; antiatherosclerotic;
 KW anorectic; antibacterial; fungicide; protozoacide; virucide; analgesic;
 KW cytostatic; immunomodulator; metabolic; antiasthmatic; antiparkinsonian;
 KW hypertensive; hypotensive; osteopathic; antiinflammatory; anti-HIV;
 KW antiinfertility; neuroprotective; antiangiinal; cardiant; antiulcer;
 KW antiallergic; nootropic; tranquiliser; neuroleptic; antidepressant;
 KW antimanic; anticonvulsant; haemostatic; immunosuppressive; hepatotropic;
 KW cerebroprotective; dermatological; antiaddictive; gynaecological; cancer;
 KW nephrotropic; gene therapy; vaccine; developmental disease; diabetes;
 KW cardiomyopathy; atherosclerosis; neurodegenerative disorder;
 KW autoimmune disorder; infectious disease; chromosome 9; gene; ss.

XX OS Homo sapiens.

XX PN WO200226985-A2.

XX XX 04-APR-2002.

XX PF 28-SEP-2001; 2001WO-US030552.

XX PR 28-SEP-2000; 2000US-0236284P.

XX PR 28-SEP-2000; 2000US-0236286P.

XX PR 03-OCT-2000; 2000US-0237581P.

XX PR 06-OCT-2000; 2000US-0238735P.

XX PR 16-OCT-2000; 2000US-0240736P.

XX PR 05-JAN-2001; 2001US-0260019P.

XX PR 08-JAN-2001; 2001US-0260338P.

XX PR 17-JAN-2001; 2001US-0262156P.

XX PR 18-JAN-2001; 2001US-0262498P.

XX PR 19-JAN-2001; 2001US-0263133P.

XX PR 24-JAN-2001; 2001US-0263691P.

XX PR 02-FEB-2001; 2001US-0266109P.

XX PR 26-FEB-2001; 2001US-0271634P.

XX PR 27-SEP-2001; 2001US-00965422.

XX (CURA-) CURAGEN CORP.

XX SPytek KA, Casman S, Padigar M, Dickson K, Vernet C;

PI Spaderna SK, Shenoy S, Gerlach V, Ellerman K, Edinger S;

PI Macdougall JR, Smithson G, Li L, Malyankar UM, Taylor S, Gunther E;

PI Tchernev VT;

XX WPI; 2002-499868/53.

DR P-PSDB; ABB81440.

XX Novel G-protein coupled-receptor polypeptides and nucleic acids for

PT diagnosing, treating cardiomyopathy, atherosclerosis, cancer,
 PT neurodegenerative, autoimmune disorders, infectious diseases and
 PT diabetes.

XX Claim 8; Page 11; 213pp; English.

XX ABN89112 to ABN89130 encoding the human G protein-coupled receptor
 CC (GPCR) proteins given in ABB81440 to ABB81458. GPCR proteins can have
 CC antiatherosclerotic, anorectic, antibacterial, fungicide, protozoacide,
 CC virucide, analgesic, cytostatic, immunomodulator, metabolic, hypotensive,
 CC antiasthmatic, antiparkinsonian, hypertensive, osteopathic, anti-HIV,
 CC antiinflammatory, antiinfertility, neuroprotective, antiangiinal,
 CC cardiant, antiulcer, antiallergic, nootropic, tranquiliser, hepatotropic,
 CC antidepressant, antimanic, anticonvulsant, haemostatic, antiaddictive,
 CC immunosuppressive, cerebroprotective, dermatological, antiaddictive,
 CC nephrotropic and gynaecological activities, and can be used in vaccines
 CC and gene therapy. GPCR proteins and polynucleotides can be used for
 CC treating or preventing a GPCR-associated disorder such as
 CC cardiomyopathy, atherosclerosis, a disorder related to cell signal
 CC processing and metabolic pathway modulation, in a human. GPCR proteins
 CC and polynucleotides can also be used in the manufacture of a medicament
 CC for preventing or treating disorders or syndromes including developmental
 CC diseases, cardiomyopathy, atherosclerosis, cancer, neurodegenerative
 CC disorders, autoimmune disorders, infectious diseases and diabetes.
 CC ABN89131 to ABN89187 represent PCR primers and probes used in the
 CC exemplification of the present invention

SQ Sequence 1329 BP; 275 A; 397 C; 326 G; 331 T; 0 U; 0 Other;

Query Match 74.4%; Score 18.6; DB 6; Length 1329;

Best Local Similarity 84.0%; Pred. No. 2.8e+02;

Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 AAAAAACACCACTGAGCTGGG 25

Db 748 AAAAAACCACTGAGCTGGG 724

RESULT 22

ABA95482

ID ABA95482 standard; cDNA; 1344 BP.

XX AC ABA95482;

XX DT 12-MAR-2002 (first entry)

XX DE Human laminin protein B2 chain 13 coding sequence.

XX KW Human; laminin protein B2 chain 13; cytostatic; immunomodulatory;
 KW antiinflammatory; neuroprotective; anti-HIV; neurological disease;
 KW phlogosis; cancer; allopasia; HIV infection; immunological disease;
 KW inflammatory disease; inflammation; ss.

OS Homo sapiens.

XX Key Location/Qualifiers

XX CDS 171..530

XX /*tag= a

XX /product= "Human laminin protein B2 chain 13"

XX WO200190165-A1.

XX PD 29-NOV-2001.

XX PF 28-APR-2001; 2001WO-CN000641.

XX PR 29-APR-2000; 2000CN-00115524.

XX (SHAN-) SHANGHAI BIONDOWN GENE DEV INC.

XX Mao Y, Xie Y;

XX WPI; 2002-106178/14.

DR P-PSDB; AAM48188.
XX HOMO laminin protein B2 chain 13 polynucleotide and polypeptide, useful
PT in diagnosis and treatment of neurological diseases, phlogosis, cancer,
XX immunological diseases and various inflammatory diseases.
XX
XX Claim 6; Page 30-31; 39pp; Chinese.
XX
XX The present sequence is the coding sequence for human laminin protein B2
CC chain 13. The protein and its coding sequence are useful in the diagnosis
CC and treatment of neurological diseases, phlogosis, cancer, allopasia,
CC HIV infection, immunological diseases and various inflammatory diseases
XX
XX Sequence 1344 BP; 380 A; 293 C; 301 G; 370 T; 0 U; 0 Other;

Query Match 74.4%; Score 18.6; DB 6; Length 1344;
Best Local Similarity 84.0%; Pred. No. 2.8e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAACACCACTGAGCTGGG 25
Db 1295 AAAAAACACCACTGAGTGAATGTG 1319

RESULT 23
ABK92484
ID ABK92484 standard; DNA; 1450 BP.
XX
XX AC ABK92484;
XX
XX 19-AUG-2002 (first entry)
XX
XX Human prostate specific nucleic acid #103.
XX
XX Prostate specific polypeptide; metastasis; prostate cancer; cancer;
KW non-cancerous prostate disease; gene therapy; gene; ds.
XX
XX Homo sapiens.
OS
XX WO200238810-A2.
XX
XX 16-MAY-2002.
XX
XX 06-NOV-2001; 2001WO-US047001.
XX
XX 06-NOV-2000; 2000US-0246109P.
XX
XX (DIAD-) DIADEXUS INC.
XX
XX Sun Y, Recipon H, Chen S, Liu C;
PI
XX WPI; 2002-471559/50.
XX
XX New prostate specific polypeptides and polynucleotides useful for
PT detecting, diagnosing, monitoring, treating, staging and predicting
PT cancers in humans having cancer and non-cancerous prostate disease.
XX
XX Claim 1; Page 203-204; 267pp; English.

XX The invention describes an isolated prostate specific polypeptide (I) and
XX nucleic acid (II) encoding it and are useful for diagnosing and
XX monitoring the presence and metastases of prostate cancer in a patient.
XX (I), (II) and an antibody to (II) are useful in quantitative and
XX qualitative diagnostic assays and methods for detecting, diagnosing,
XX monitoring, treating, staging and predicting cancer in humans having
XX cancer or may have the risk of developing cancer. (I) and (II) are also
XX useful: for determining non-cancerous prostate disease, by measuring
XX their expression levels and/or structural alterations; for determining
XX the sample that has prostate tissue-like characteristics or is a prostate
XX tissue; as an element in an array or a multi-analyte test to recognise
XX expression patterns associated with prostate cancer and other prostate
XX related disorders; and as elements in a computer program for pattern
XX recognition of prostate disorders. (I) and (II) are useful for producing

CC engineered prostate tissue for treatment and research. (II) is useful for
CC producing transgenic animals and cells and also in gene therapy. This
CC sequence represents a prostate specific nucleic acid described in the
CC invention
XX
XX Sequence 1450 BP; 341 A; 370 C; 418 G; 321 T; 0 U; 0 Other;

Query Match 74.4%; Score 18.6; DB 6; Length 1450;
Best Local Similarity 84.0%; Pred. No. 2.8e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAACACCACTGAGCTGGG 25
Db 619 AAAAAACCACTGTGTGGTG 643

RESULT 24
ABK95452/c
ID ABK95452 standard; DNA; 2027 BP.

XX
XX AC ABK95452;

XX
XX 24-SEP-2002 (first entry)

XX
XX DNA encoding novel G-protein coupled receptor related protein #11.

XX
XX G protein coupled receptor; GPCR; olfactory receptor;
KW cell signal processing disorder; metabolic pathway modulation;
KW cardiomyopathy; atherosclerosis; diabetes; developmental disease;
KW immune disease; taste disorder; scent detectability disorder; obesity;
KW Burkitt's lymphoma; corticosteroid disease; infectious disease; pain;
KW signal transduction pathway disorder; metabolic pathway disorder;
KW retinal disease; metabolic disorder; cancer; Parkinson's disease;
KW acute heart failure; urinary retention; osteoporosis; Crohn's disease;
KW ulcer; allergy; neurological disorder; genetic disorder; transplantation;
KW fertility; pancreatitis; hyperthyroidism; Endometriosis;
KW forensic biology; transgenic animal; gene; ds.

XX
XX Mus musculus.

XX
XX WO200240539-A2.

XX
XX 23-MAY-2002.

XX
XX 16-OCT-2001; 2001WO-US032256.

XX
XX 16-OCT-2000; 2000US-0240704P.

XX
XX 26-OCT-2000; 2000US-0243497P.

XX
XX 31-OCT-2000; 2000US-024542P.

XX
XX 03-NOV-2000; 2000US-0245484P.

XX
XX 12-DEC-2000; 2000US-0255017P.

XX
XX 17-JAN-2001; 2001US-0262159P.

XX
XX 22-JAN-2001; 2001US-0263216P.

XX
XX 25-JAN-2001; 2001US-0263340P.

XX
XX 12-FEB-2001; 2001US-0264118P.

XX
XX 15-FEB-2001; 2001US-0268225P.

XX
XX 27-JUL-2001; 2001US-0269031P.

XX
XX 27-JUL-2001; 2001US-0308203P.

XX
XX (CURA-) CURAGEN CORP.

XX
XX Kekuda R, Spytek KA, Casman SJ, Zethusen BD, Li L, Tchernev VT;

XX
XX Colman SD, Ballinger RA, Padigar M, Wolenc AR, Shenoy SG;

XX
XX Edinger SR, Gerlach V, Gangoli EA, Macdougall JR, Smithson G;

XX
XX Peyman JA, Stone DJ, Gunther E, Ellerman K, Grosse WM, Alsobrook JP;

XX
XX Lepley DM, Burgess CB;
XX
XX WPI; 2002-500205/53.
XX
XX P-PSDB; ABG66934.
XX
XX Novel G protein coupled receptor especially olfactory receptor
PT polypeptides and nucleic acids for diagnosing and treating
PT atherosclerosis, cardiomyopathy and diabetes.

```

XX Claim 8; Page 43; 309pp; English.
PS
XX
CC The invention describes an isolated G protein coupled receptor X (GPCR1-
CC 12) polypeptide, especially an olfactory receptor. GPCRX polypeptides are
CC useful for identifying an agent that binds to the polypeptide and for
CC identifying a candidate substance or ligand molecules interacting with an
CC olfactory receptor polypeptide. The polypeptide, (I) and (II) are also
CC useful for treating diseases and disorders related to cell signal
CC processing and metabolic pathway modulation e.g. cardiomyopathy,
CC atherosclerosis and diabetes, and developmental diseases, immune
CC diseases, taste and scent detectability disorders, Burkitt's lymphoma,
CC corticosterogenic disease, signal transduction pathway disorders,
CC metabolic pathway disorders, retinal diseases, metabolic disorders,
CC obesity, infectious disease, pain, cancer, Parkinson's disease, acute
CC heart failure, urinary retention, osteoporosis, Crohn's disease, ulcers,
CC allergies, neurological disorders, genetic disorders, transplantation,
CC fertility, Pancreatitis, Hyperthyroidism and Endometriosis. GPCRX
CC sequences are also useful for identifying a cell or tissue type in a
CC biological sample, to amplify DNA sequences from very small biological
CC samples such as tissues e.g. hair or skin or body fluids in forensic
CC biology. Cells comprising (I) are useful for producing non-human
CC transgenic animals for studying the function and/or activity of GPCRX
CC protein and for identifying and/or evaluating modulators of GPCRX protein
CC activity. This sequence encodes a novel G-protein coupled receptor
CC described in the invention.
XX
SQ Sequence 2027 BP; 480 A; 548 C; 505 G; 494 T; 0 U; 0 Other;

Query Match          74.4%; Score 18.6; DB 6; Length 2027;
Best Local Similarity 84.0%; Pred. No. 3e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAACACCACCACTGAGCTGGG 25
   ||||| ||||| ||||| |||||
Db 1215 AAAAAACACCACCACTGAGCTGGT 1191

RESULT 25
ABK40203/c
ID ABK40203 standard; cDNA; 2028 BP.
XX
AC ABK40203;
XX
DT 21-MAY-2002 (first entry)
XX
DE Human G protein-coupled receptor (GPCR) 9 cDNA #1.
XX
KW Human; G protein-coupled receptor; GPCR; cell signalling processing; ss;
KW cardiomyopathy; atherosclerosis; developmental disease; immune disease;
KW non insulin-dependent diabetes mellitus; NIDDM1; bacterial infection;
KW fungal infection; protozoal infection; viral infection; adenocarcinoma;
KW human immunodeficiency virus; HIV; cancer; lymphoma; anorexia; asthma;
KW neurodegenerative disorder; Parkinson's disease; acute heart failure;
KW hypotension; hypertension; osteoporosis; Crohn's disease; allergy; ulcer;
KW multiple sclerosis; osteodystrophy; angina pectoris; schizophrenia; gene;
KW myocardial infarction; neurological disorder; anxiety; manic depression;
KW Tourette's syndrome; Huntington's disease; Alzheimer's disease; obesity;
XX Homo sapiens.
XX WO200202637-A2.
PN
XX
XX 10-JAN-2002.
PD
XX
XX 02-JUL-2001; 2001WO-US021174.
PF
XX
XX 30-JUN-2000; 2000US-0215483P.
PR
XX 30-JUN-2000; 2000US-0215495P.
PR
XX 06-JUL-2000; 2000US-0216117P.
PR
XX 25-JUL-2000; 2000US-0220589P.
PR
XX 02-NOV-2000; 2000US-0245294P.
PR

```

```

PR 10-JAN-2001; 2001US-0260729P.
PR 10-JAN-2001; 2001US-0260851P.
PR 26-FEB-2001; 2001US-0271673P.
XX
XX (CURA-) CURAGEN CORP.
XX
XX Spaderna SK, Padigaru M, Spytek KA, Casman S, Rastelli L;
PI Mayankar U, Tchernev V;
XX
XX WPI; 2002-164524/21.
DR P-PSDB; AAU85873.
XX
XX Novel polypeptides and nucleic acids of G protein coupled receptor useful
XX for diagnosing and treating diseases e.g. anorexia, asthma, Parkinson's
XX disease, acute heart failure and infections.
XX
XX Claim 9; Page 56-57; 202pp; English.
XX
XX The invention relates to human G protein coupled receptor (GPCR)
XX polypeptides and the polynucleotides encoding them. GPCR polypeptides are
XX useful for treating or preventing a GPCR-associated disorder which is
XX related to cell signalling processing and metabolic pathway modulation,
XX which includes cardiomyopathy and atherosclerosis in a human. The
XX sequences are useful in treatment of patients suffering from
XX developmental diseases, immune diseases, non insulin-dependent diabetes
XX mellitus (NIDDM1), bacterial, fungal, protozoal and viral infections
XX (e.g. infections caused by HIV), cancer (e.g. adenocarcinoma, lymphoma),
XX anorexia, asthma, neurodegenerative disorders (e.g. Parkinson's disease),
XX acute heart failure, hypotension, hypertension, osteoporosis, Crohn's
XX disease, multiple sclerosis, osteodystrophy, angina pectoris, myocardial
XX infarction, ulcers, allergies, neurological disorders including anxiety,
XX schizophrenia, manic depression, delirium, dementia, Huntington's
XX disease, Alzheimer's disease, Tourette's syndrome, metabolic disorders
XX and obesity. This sequence represents cDNA encoding a GPCR polypeptide of
XX the invention
XX
SQ Sequence 2028 BP; 480 A; 549 C; 505 G; 494 T; 0 U; 0 Other;

Query Match          74.4%; Score 18.6; DB 6; Length 2028;
Best Local Similarity 84.0%; Pred. No. 3e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAACACCACCACTGAGCTGGG 25
   ||||| ||||| ||||| |||||
Db 1216 AAAAAACACCACCACTGAGCTGGT 1192

RESULT 26
ABZ58229/c
ID ABZ58229 standard; cDNA; 2312 BP.
XX
AC ABZ58229;
XX
XX 28-APR-2003 (first entry)
XX
XX Human G-protein coupled receptor GCRC-6 (olfactory receptor) cDNA.
XX
XX GCRC-6; G-protein coupled receptor; olfactory receptor; receptor; human;
XX vasotropic; cardiac; cardiovascular; antiarteriosclerotic; antianginal;
XX antiinflammatory; cerebroprotective; antiemetic; antidiabetic; antigout;
XX neuroprotective; nootropic; tranquilizer; osteopathic; cytostatic;
XX hypotensive; gastrointestinal; antiulcer; antinaemic; antithyroid;
XX anticonvulsant; antiparkinsonian; anorectic; virucide; olfaction; taste;
XX gene therapy; gene; ss.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 1371..2288
XX /*tag= a
XX /*product= "Human GCRC-6"
XX
XX WO2003000859-A2.
PN

```

```

XX 03-JAN-2003.
PD
XX
PF 20-JUN-2002; 2002WO-US020036.
XX
XX 22-JUN-2001; 2001US-0300494P.
PR 13-JUL-2001; 2001US-0305326P.
PR 27-JUL-2001; 2001US-0308165P.
PR 03-AUG-2001; 2001US-0310115P.
PR 10-AUG-2001; 2001US-0311413P.
PR 24-AUG-2001; 2001US-0314679P.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Baughn MR, Walia NK, Richardson TW, Ison CH, Yang J, Emerling BM;
PI Thornton M, Borowsky ML, Ramkumar J, Yue H, Graul RC, Hafalia AJA,
PI Lu DAM, Lu Y, Kallick DA, Gandhi AR, Warren BA, Lee EA;
PI Swarnakar A, Walsh RT, Yao MG, Griffin JA, Au-Young JK, Khan FA;
PI Lee SY, Elliott VS;
XX
XX WPI; 2003-221414/21.
DR P-PSDB; ABP72212.
XX
XX New human G-protein coupled receptor proteins and polynucleotides, useful
PT for diagnosing, treating or preventing e.g. ischemic heart disease,
PT Crohn's disease, AIDS, stroke, Creutzfeldt-Jakob disease, leukemia or
PT lymphoma.
XX
XX Claim 5; Page 171-172; 182pp; English.
XX
XX The present sequence is that of a cDNA clone, denoted Incyte
CC Polynucleotide ID No. 55153201CB1, encoding novel human G-protein coupled
CC receptor GCRC-6. The encoded protein is an olfactory receptor, being 73%
CC identical from residue 1 to 296 to a murine olfactory receptor, and
CC including a 7-transmembrane receptor (rhodopsin family) domain. The
CC invention provides GCRC-1 to -26 polynucleotides (see AB258224-49),
CC polypeptides (see ABP72207-32) and antibodies. These are useful for the
CC diagnosis, treatment or prevention of neural disorders of olfaction, cell
CC proliferative, neurological, cardiovascular, gastrointestinal,
CC autoimmune, inflammatory, and metabolic disorders, and viral infections.
CC They are also used in claimed methods of screening for agonist and
CC antagonist compounds, in treating a disease or condition associated with
CC decreased expression or overexpression of a functional GCRC, in
CC assessing the toxicity of a test compound, in the construction of a
CC microarray and generation of an expression profile, and in the
CC identification of a compound which modulates, mimics and/or blocks an
CC olfactory and/or taste sensation
XX
XX Sequence 2312 BP; 564 A; 636 C; 568 G; 544 T; 0 U; 0 Other;
SQ
Query Match 74.4%; Score 18.6; DB 8; Length 2312;
Best Local Similarity 84.0%; Pred. No. 3e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 AAAAAACACACCACTGACCTGGG 25
Db 2115 AAAAAACACCACTGCGGTG 2091
RESULT 27
AAK68477
ID AAK68477 standard; DNA; 2464 BP.
XX
XX AAK68477;
AC
XX
XX 06-NOV-2001 (first entry)
DT
XX
XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:23289.
DE
XX
XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX
XX Homo sapiens.
OS

```

```

XX WO200157182-A2.
PN
XX
PD
XX
XX 09-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US001354.
XX
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-022547P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.

```

PR 27-SEP-2000; 2000US-0235836P.
 PR 29-SEP-2000; 2000US-0236327P.
 PR 29-SEP-2000; 2000US-0236367P.
 PR 29-SEP-2000; 2000US-0236368P.
 PR 29-SEP-2000; 2000US-0236369P.
 PR 29-SEP-2000; 2000US-0236370P.
 PR 02-OCT-2000; 2000US-0236802P.
 PR 02-OCT-2000; 2000US-0237037P.
 PR 02-OCT-2000; 2000US-0237038P.
 PR 02-OCT-2000; 2000US-0237039P.
 PR 02-OCT-2000; 2000US-0237040P.
 PR 13-OCT-2000; 2000US-0239935P.
 PR 13-OCT-2000; 2000US-0239937P.
 PR 20-OCT-2000; 2000US-0240960P.
 PR 20-OCT-2000; 2000US-0241221P.
 PR 20-OCT-2000; 2000US-0241785P.
 PR 20-OCT-2000; 2000US-0241786P.
 PR 20-OCT-2000; 2000US-0241787P.
 PR 20-OCT-2000; 2000US-0241808P.
 PR 20-OCT-2000; 2000US-0241809P.
 PR 20-OCT-2000; 2000US-0241826P.
 PR 01-NOV-2000; 2000US-0244617P.
 PR 08-NOV-2000; 2000US-0244647P.
 PR 08-NOV-2000; 2000US-0246475P.
 PR 08-NOV-2000; 2000US-0246476P.
 PR 08-NOV-2000; 2000US-0246477P.
 PR 08-NOV-2000; 2000US-0246478P.
 PR 08-NOV-2000; 2000US-0246523P.
 PR 08-NOV-2000; 2000US-0246524P.
 PR 08-NOV-2000; 2000US-0246525P.
 PR 08-NOV-2000; 2000US-0246526P.
 PR 08-NOV-2000; 2000US-0246527P.
 PR 08-NOV-2000; 2000US-0246528P.
 PR 08-NOV-2000; 2000US-0246532P.
 PR 08-NOV-2000; 2000US-0246609P.
 PR 08-NOV-2000; 2000US-0246610P.
 PR 08-NOV-2000; 2000US-0246611P.
 PR 08-NOV-2000; 2000US-0246613P.
 PR 17-NOV-2000; 2000US-0249207P.
 PR 17-NOV-2000; 2000US-0249208P.
 PR 17-NOV-2000; 2000US-0249209P.
 PR 17-NOV-2000; 2000US-0249210P.
 PR 17-NOV-2000; 2000US-0249211P.
 PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249244P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249264P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 05-DEC-2000; 2000US-0256719P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 06-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 11-DEC-2000; 2000US-0251990P.
 PR 05-JAN-2001; 2001US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.
 (HUMA-) HUMAN GENOME SCI INC.
 XX
 PA
 XX
 PI Rosen CA, Barash SC, Ruben SM;
 XX WPI; 2001-483426/52.
 XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides;
 PT useful for preventing, diagnosing and/or treating cancers and metastasis.
 XX Disclosure; SEQ ID NO 23289; 3071pp + Sequence Listing; English.
 XX AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
 CC amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic
 CC activity, and can be used in gene therapy and vaccine production. (I)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patients own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I), by inserting the
 CC nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/haematopoietic-related diseases, especially
 CC cancers and cancer metastases of haematopoietic-derived cells. AAK4703
 CC to AAK87694 represent human immune/haematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
 CC represent sequences used in the exemplification of the present invention
 XX
 SQ Sequence 2464 BP; 529 A; 658 C; 791 G; 486 T; 0 U; 0 Other;
 Query Match 74.4%; Score 18.6; DB 4; Length 2464;
 Best Local Similarity 84.0%; Pred. No. 3.1e+02;
 Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 1 AAAAAACACCACTGAGCTGGG 25
 DB 1751 AAAAAACCAACAAACTGAGCTGGG 1775
 RESULT 28
 AAK68479
 ID AAK68479 standard; DNA; 2464 BP.
 XX
 AC AAK68479;
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:23291.
 XX
 KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
 KW cytostatic; gene therapy; vaccine; metastasis; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200157182-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 17-JAN-2001; 2001WO-US001354.
 XX
 XX 31-JAN-2000; 2000US-0179065P.
 PR 04-FEB-2000; 2000US-0180628P.
 PR 24-FEB-2000; 2000US-0184664P.
 PR 02-MAR-2000; 2000US-0186350P.
 PR 16-MAR-2000; 2000US-0189874P.
 PR 17-MAR-2000; 2000US-0190076P.
 PR 18-APR-2000; 2000US-0198123P.
 PR 19-MAY-2000; 2000US-0205515P.
 PR 07-JUN-2000; 2000US-0209467P.
 PR 28-JUN-2000; 2000US-0214886P.
 PR 30-JUN-2000; 2000US-0215135P.
 PR 07-JUL-2000; 2000US-0216647P.
 PR 11-JUL-2000; 2000US-0216880P.
 PR 17-JUL-2000; 2000US-0217487P.

PR 11-JUL-2000; 2000US-0217496P.
 PR 14-JUL-2000; 2000US-0218290P.
 PR 26-JUL-2000; 2000US-0220863P.
 PR 14-AUG-2000; 2000US-0220964P.
 PR 14-AUG-2000; 2000US-0224518P.
 PR 14-AUG-2000; 2000US-0224519P.
 PR 14-AUG-2000; 2000US-0225213P.
 PR 14-AUG-2000; 2000US-0225214P.
 PR 14-AUG-2000; 2000US-0225266P.
 PR 14-AUG-2000; 2000US-0225267P.
 PR 14-AUG-2000; 2000US-0225268P.
 PR 14-AUG-2000; 2000US-0225270P.
 PR 14-AUG-2000; 2000US-0225447P.
 PR 14-AUG-2000; 2000US-0225757P.
 PR 14-AUG-2000; 2000US-0225758P.
 PR 14-AUG-2000; 2000US-0225759P.
 PR 18-AUG-2000; 2000US-0226279P.
 PR 22-AUG-2000; 2000US-0226681P.
 PR 22-AUG-2000; 2000US-0226868P.
 PR 22-AUG-2000; 2000US-0227182P.
 PR 23-AUG-2000; 2000US-0227009P.
 PR 30-AUG-2000; 2000US-0228924P.
 PR 01-SEP-2000; 2000US-0229287P.
 PR 01-SEP-2000; 2000US-0229343P.
 PR 01-SEP-2000; 2000US-0229344P.
 PR 01-SEP-2000; 2000US-0229345P.
 PR 05-SEP-2000; 2000US-0229509P.
 PR 05-SEP-2000; 2000US-0229513P.
 PR 06-SEP-2000; 2000US-0230437P.
 PR 06-SEP-2000; 2000US-0230438P.
 PR 08-SEP-2000; 2000US-0231242P.
 PR 08-SEP-2000; 2000US-0231243P.
 PR 08-SEP-2000; 2000US-0231244P.
 PR 08-SEP-2000; 2000US-0231413P.
 PR 08-SEP-2000; 2000US-0231414P.
 PR 08-SEP-2000; 2000US-0232080P.
 PR 08-SEP-2000; 2000US-0232081P.
 PR 12-SEP-2000; 2000US-0231968P.
 PR 14-SEP-2000; 2000US-0232397P.
 PR 14-SEP-2000; 2000US-0232398P.
 PR 14-SEP-2000; 2000US-0232399P.
 PR 14-SEP-2000; 2000US-0232400P.
 PR 14-SEP-2000; 2000US-0232401P.
 PR 14-SEP-2000; 2000US-0233063P.
 PR 14-SEP-2000; 2000US-0233064P.
 PR 14-SEP-2000; 2000US-0233065P.
 PR 21-SEP-2000; 2000US-0234223P.
 PR 21-SEP-2000; 2000US-0234274P.
 PR 25-SEP-2000; 2000US-0234937P.
 PR 25-SEP-2000; 2000US-0234938P.
 PR 26-SEP-2000; 2000US-0235484P.
 PR 27-SEP-2000; 2000US-0235834P.
 PR 27-SEP-2000; 2000US-0235836P.
 PR 29-SEP-2000; 2000US-0236327P.
 PR 29-SEP-2000; 2000US-0236367P.
 PR 29-SEP-2000; 2000US-0236368P.
 PR 29-SEP-2000; 2000US-0236369P.
 PR 29-SEP-2000; 2000US-0236370P.
 PR 02-OCT-2000; 2000US-0236802P.
 PR 02-OCT-2000; 2000US-0237037P.
 PR 02-OCT-2000; 2000US-0237038P.
 PR 02-OCT-2000; 2000US-0237039P.
 PR 02-OCT-2000; 2000US-0237040P.
 PR 13-OCT-2000; 2000US-0239935P.
 PR 13-OCT-2000; 2000US-0239937P.
 PR 20-OCT-2000; 2000US-0240960P.
 PR 20-OCT-2000; 2000US-0241221P.
 PR 20-OCT-2000; 2000US-0241785P.
 PR 20-OCT-2000; 2000US-0241786P.
 PR 20-OCT-2000; 2000US-0241787P.
 PR 20-OCT-2000; 2000US-0241808P.
 PR 20-OCT-2000; 2000US-0241809P.
 PR 20-OCT-2000; 2000US-0241826P.

PR 01-NOV-2000; 2000US-0244617P.
 PR 08-NOV-2000; 2000US-0246474P.
 PR 08-NOV-2000; 2000US-0246475P.
 PR 08-NOV-2000; 2000US-0246476P.
 PR 08-NOV-2000; 2000US-0246477P.
 PR 08-NOV-2000; 2000US-0246478P.
 PR 08-NOV-2000; 2000US-0246523P.
 PR 08-NOV-2000; 2000US-0246524P.
 PR 08-NOV-2000; 2000US-0246525P.
 PR 08-NOV-2000; 2000US-0246526P.
 PR 08-NOV-2000; 2000US-0246527P.
 PR 08-NOV-2000; 2000US-0246528P.
 PR 08-NOV-2000; 2000US-0246532P.
 PR 08-NOV-2000; 2000US-0246609P.
 PR 08-NOV-2000; 2000US-0246610P.
 PR 08-NOV-2000; 2000US-0246611P.
 PR 08-NOV-2000; 2000US-0246613P.
 PR 17-NOV-2000; 2000US-0249207P.
 PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249244P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249264P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249279P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251031P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 05-DEC-2000; 2000US-0256719P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 08-DEC-2000; 2000US-0251990P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-483426/52.

Nucleic acids encoding human immune/hematopoietic antigen polypeptides, useful for preventing, diagnosing and/or treating cancers and metastasis.

Disclosure; SEQ ID NO 23291; 3071pp + Sequence Listing; English.

AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I) amino acid sequences given in AA82170 to AA82191. (I) have cytostatic activity, and can be used in gene therapy and vaccine production. (I) proteins and polynucleotides may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate (I) expression. For example, they may be used to treat disorders associated with decreased expression by rectifying mutations or deletions in a patient's genome that affect the activity of (I) by expressing inactive proteins or to supplement the patient's own production of (I). Additionally, (I) polynucleotides may be used to produce the secreted (I), by inserting the nucleic acids into a host cell and culturing the cell to express the protein. (I) proteins and polynucleotides may be used to prevent,

CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
CC represent sequences used in the exemplification of the present invention
XX
SQ Sequence 2464 BP; 529 A; 661 C; 790 G; 484 T; 0 U; 0 Other;

Query Match 74.4%; Score 18.6; DB 4; Length 2464;
Best Local Similarity 84.0%; Pred. No. 3.1e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAACACCACTGAGCTGGG 25
||||||| | ||||| |||||
Db 1751 AAAAAACACCAAACTGACTGGG 1775

RESULT 29
AAL04040/c
ID AAL04040 standard; DNA; 2464 BP.
XX
AC AAL04040;
XX
DT 21-NOV-2001 (first entry)
XX
DE Human reproductive system related antigen DNA SEQ ID NO: 6728.
XX
KW Human; reproductive system related antigen; reproductive system disorder;
KW cancer; gene therapy; ds.
XX
OS Homo sapiens.
XX
PN WO200155320-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001339.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184666P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.

PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.

```
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251980P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-02559678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Baraah SC, Ruben SM;
XX
DR WPI; 2001-465570/50.
XX
PT Isolated nucleic acid molecule encoding a reproductive system antigen is
PT used in preventing, treating or ameliorating a medical condition.
XX
XX
PS Disclosure; SEQ ID NO 6728; 1297pp + Sequence Listing; English.
XX
CC The present invention provides the protein and coding sequences of a
CC number of human reproductive system related antigens. These can be used
CC in the prevention and treatment of reproductive system disorders,
CC including cancer. The present sequence is a genomic sequence encoding a
CC protein of the invention
XX
SQ Sequence 2464 BP; 486 A; 791 C; 658 G; 529 T; 0 U; 0 Other;
Query Match 74.4%; Score 18.6; DB 4; Length 2464;
Best Local Similarity 84.0%; Pred. No. 3.1e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Qy 1 AAAAAACAACCACTGAGCTGGG 25
Db 714 AAAAAACAACCACTGAGCTGGG 690
RESULT 30
AAL04041/c
ID AAL04041 standard; DNA; 2464 BP.
XX
AC AAL04041;
XX
DT 21-NOV-2001 (first entry)
XX
DE Human reproductive system related antigen DNA SEQ ID NO: 6729.
XX
KW Human; reproductive system related antigen; reproductive system disorder;
KW cancer; gene therapy; ds.
XX
OS Homo sapiens.
XX
FN WO200155320-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001339.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 16-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 14-JUL-2000; 2000US-0217496P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 23-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 29-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
```


CC nucleic acid-associated proteins (NAAP). The DNA and protein sequences of
CC the invention are useful in diagnosing, preventing and treating
CC diseases/conditions associated with altered expression of NAAP, such as:
CC autoimmune/inflammatory disorders (e.g. AIDS and allergies), infections
CC (e.g. bacterial and viral), metabolic disorders (e.g. obesity),
CC reproductive disorders (e.g. infertility), neurological disorders (e.g.
CC Parkinson's disease and Alzheimer's disease), cardiovascular disorders
CC (e.g. myocardial infarction and hypertension), eye disorders, or cell
CC proliferative diseases (e.g. cancer). The present DNA sequence encodes a
CC human NAAP protein of the invention.

XX
SQ Sequence 2775 BP; 845 A; 584 C; 738 G; 608 T; 0 U; 0 Other;

Query Match 74.4%; Score 18.6; DB 12; Length 2775;
Best Local Similarity 84.0%; Pred. No. 3.1e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAACACACCACCTGAGCTGGG 25
|||||
Db 2705 AAAAAACACACCACCTGGAATGTG 2729

RESULT 32

ID ADQ64688 standard; cDNA; 3009 BP.

XX AC ADQ64688;

XX DT 07-OCT-2004 (first entry)

XX DE Novel human cDNA sequence #1849.

XX ss; gene; osteopathic; neuroprotective; nootropic; antiparkinsonian;
KW cytotatic; gene therapy; diagnostic marker; morbid state; osteoporosis;
KW neurological disease; Alzheimer's disease; Parkinson's disease; dementia;
KW cancer.

XX OS Homo sapiens.

XX PN EP1440981-A2.

XX PD 28-JUL-2004.

XX PF 21-JAN-2004; 2004EP-00001196.

XX PR 21-JAN-2003; 2003JP-00102206.

XX PR 05-MAY-2003; 2003JP-00131392.

XX PA (REAS-) RES ASSOC BIOTECHNOLOGY.

XX PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;

PI Yamamoto J, Isono Y, Nagai K, Irie R;

XX WPI; 2004-535376/52.

XX P-PSDB; ADQ66876.

XX Novel 2495 cDNA, useful for treating osteoporosis, neurological diseases,
PT Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.

XX Claim 1; SEQ ID NO 1849; 2449pp; English.

XX The invention relates to 2495 novel polynucleotides (I) and their encoded
CC polypeptides, sequences hybridizing to these nucleotides, sequences
CC encoding partial polypeptides and sequences having 70% or 90% identity to
CC the nucleotide and protein sequences. The nucleotides and polypeptides
CC are useful as diagnostic markers or therapeutic target for the diseases
CC or morbid states. They are also useful for treating osteoporosis,
CC neurological diseases, Alzheimer's diseases, Parkinson's diseases,
CC dementia and various cancers. This sequence corresponds to a nucleotide
XX sequence of the invention.

XX Sequence 3009 BP; 819 A; 626 C; 732 G; 832 T; 0 U; 0 Other;

Query Match 74.4%; Score 18.6; DB 12; Length 3009;
Best Local Similarity 84.0%; Pred. No. 3.2e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAACACACCACCTGAGCTGGG 25
|||||
Db 2955 AAAAAACACACCACCTGGAATGTG 2979

RESULT 33

ID ADQ24417 standard; DNA; 3029 BP.

XX AC ADQ24417;

XX DT 26-AUG-2004 (first entry)

XX DE Human soft tissue sarcoma-upregulated DNA - SEQ ID 7237.

XX soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human;
KW ds.

XX OS Homo sapiens.

XX PN WO2004048938-A2.

XX PD 10-JUN-2004.

XX PF 26-NOV-2003; 2003WO-US038193.

XX PR 26-NOV-2002; 2002US-0429739P.

XX PA (PROT-) PROTEIN DESIGN LABS INC.

XX PI Aziz N, Ginsburg WM, Zlotnik A;

XX WPI; 2004-441208/41.

XX Early detection of soft tissue sarcoma comprises determining expression
PT of a gene in a first soft tissue sample and a normal soft tissue sample
PT and comparing the gene expression, also useful in treating soft tissue
PT sarcoma.

XX Example 2; SEQ ID NO 7237; 210pp; English.

XX The invention relates to a novel method for detecting soft tissue sarcoma
CC which comprises obtaining a first soft tissue sample from an individual
CC and a normal soft tissue sample from the same or different individual,
CC determining the expression of a gene in both samples and comparing the
CC expression of the gene in both soft tissue samples, where a higher level
CC of protein expression in the first soft tissue sample indicates the
CC presence of soft tissue sarcoma. The method of the invention has
CC cytotatic applications and may be useful for detecting soft tissue
CC sarcoma, possibly via gene therapy or vaccine production. The nucleic
CC acid sequences may be useful in diagnostic and screening applications.
CC The current sequence is that of a human soft tissue sarcoma-upregulated
CC DNA of the invention. The current sequence is not shown within the
CC specification per se but was submitted in CD format by the inventor.

XX SQ Sequence 3029 BP; 908 A; 664 C; 774 G; 650 T; 0 U; 33 Other;

Query Match 74.4%; Score 18.6; DB 12; Length 3029;
Best Local Similarity 84.0%; Pred. No. 3.2e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAACACACCACCTGAGCTGGG 25
|||||
Db 2946 AAAAAACACACCACCTGGAATGTG 2970

RESULT 34

ID ADI16310

ID ADI16310 standard; DNA; 3030 BP.

```
XX AC ADI16310;
XX DT 22-APR-2004 (first entry)
XX DE Human nucleic acid-associated protein (NAAP) coding sequence #45.
XX KW human; nucleic acid-associated protein; NAAP; autoimmune disorder;
XX inflammatory disorder; AIDS; allergy; infection; metabolic disorder;
XX obesity; reproductive disorder; infertility; neurological disorder;
XX Parkinson's disease; Alzheimer's disease; cardiovascular disorder;
XX myocardial infarction; hypertension; eye disorder;
XX cell proliferative disease; cancer; ds; gene.
XX OS Homo sapiens.
XX PN WO2003094848-A2.
XX PD 20-NOV-2003.
XX PF 09-MAY-2003; 2003WO-US014450.
XX PR 10-MAY-2002; 2002US-0379843P.
XX PR 24-MAY-2002; 2002US-0383457P.
XX PR 31-MAY-2002; 2002US-0384699P.
XX PR 06-JUN-2002; 2002US-0387265P.
XX PA (INCY-) INCYTE CORP.
XX PI Kable AE, Elliott VS, Tran UK, Ramkumar J, Marquis JP, Chawla NK;
XX Richardson TW, Bulloch SA, Khare R, Lee SY, Lal PG, Tang YT, Yue H;
XX Swarnakar A, Becha SD, Hafalia AUA, Chang H, Baughn MR, Borowsky ML;
XX Gietzen KJ, He A, Forsythe IU, Sprague WW, Blake JJ, Warren BA;
XX Mason PM, Ison CH, Lindquist EA, Wilson AD, Jin P;
XX WPI: 2004-011999/01.
XX DR P-PSDB; ADI16260.
XX PT New human nucleic acid associated proteins and polynucleotides, useful
XX for diagnosing, preventing or treating diseases or conditions associated
XX with aberrant protein expression, e.g. cancer, AIDS, atherosclerosis or
XX stroke.
XX PS Claim 5; SEQ ID NO 95; 400pp; English.
XX PN The invention comprises the amino acid and coding sequences of human
XX nucleic acid-associated proteins (NAAP). The DNA and protein sequences of
XX the invention are useful in diagnosing, preventing and treating
XX diseases/conditions associated with altered expression of NAAP, such as:
XX autoimmune/inflammatory disorders (e.g. AIDS and allergies), infections
XX (e.g. bacterial and viral), metabolic disorders (e.g. obesity),
XX reproductive disorders (e.g. infertility), neurological disorders (e.g.
XX Parkinson's disease and Alzheimer's disease), cardiovascular disorders
XX (e.g. myocardial infarction and hypertension), eye disorders, or cell
XX proliferative diseases (e.g. cancer). The present DNA sequence encodes a
XX human NAAP protein of the invention.
XX SQ Sequence 3030 BP; 865 A; 705 C; 787 G; 673 T; 0 U; 0 Other;

Query Match 74.4%; Score 18.6; DB 12; Length 3030;
Best Local Similarity 84.0%; Pred. No. 3.2e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 AAAAAACACCACTGAGCTGGG 25
Db 2988 AAAAAACACCACTGAGCTGGG 3012

RESULT 35
ADR25447
ID ADR25447 standard; DNA; 3154 BP.
XX ADR25447;
AC
```

```
XX DT 21-OCT-2004 (first entry)
XX DE Breast cancer prognosis marker #1308.
XX KW ds; breast cancer; prognosis; gene expression; diagnosis.
XX OS Homo sapiens.
XX PN WO2004065545-A2.
XX PD 05-AUG-2004.
XX PF 15-JAN-2004; 2004WO-US001100.
XX PR 15-JAN-2003; 2003US-00342887.
XX PA (ROSE-) ROSETTA INPHARMATICS LLC.
XX PA (NECA-) NETHERLANDS CANCER INST.
XX PI Van't Veer LJ, He Y;
XX WPI: 2004-593473/57.
XX DR Classifying a breast cancer patient according to prognosis comprises
XX PT determining the similarity between the level of expression of each of
XX PT five genes in a cell sample taken from patient, to control levels.
XX PS Disclosure; SEQ ID NO 1308; 226pp; English.
XX CC The invention relates to a method of classifying a breast cancer patient
XX CC according to prognosis by determining the similarity between the level of
XX CC expression of each of five genes for which markers are listed in the
XX CC specification, in a cell sample taken from the breast cancer patient, to
XX CC control levels of expression for each respective five genes to obtain a
XX CC patient similarity value. The methods are useful for classifying a breast
XX CC cancer patient according to prognosis. Kits and computer program products
XX CC are useful for data analysis using the diagnostic, prognostic and
XX CC statistical methods of the invention. This sequence corresponds to a
XX CC marker used in the method of the invention.
XX SQ Sequence 3154 BP; 901 A; 734 C; 811 G; 708 T; 0 U; 0 Other;

Query Match 74.4%; Score 18.6; DB 13; Length 3154;
Best Local Similarity 84.0%; Pred. No. 3.2e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 AAAAAACACCACTGAGCTGGG 25
Db 3112 AAAAAACACCACTGAGCTGGG 3136

RESULT 36
AAZ57865
ID AAZ57865 standard; cDNA; 3190 BP.
XX AC AAZ57865;
XX DT 11-APR-2000 (first entry)
XX DE Protein regulating gene expression PRGE-27 cDNA clone 1399169.
XX KW Protein regulating gene expression; PRGE-27; human; cell proliferation;
XX KW antiproliferative; inflammation; antinflammatory; gene therapy;
XX KW diagnosis; ss.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX CDS 14..2869
XX FT /*tag= a
XX PN WO9964596-A2.
```

```
XX PD 16-DEC-1999.
XX PF 11-JUN-1999; 99WO-US013281.
XX PR 12-JUN-1998; 98US-0089029P.
XX PR 29-JUL-1998; 98US-0094575P.
XX PR 14-OCT-1998; 98US-0104624P.
XX PA (INCY-) INCYTE PHARM INC.
XX PI Lal P, Yue H, Tang YT, Hillman JL, Bandman O, Corley NC;
XX PI Guegler KJ, Gorgone GA, Baughn MR, Patterson C, Lu DAM;
XX PR WPI; 2000-116543/10.
XX PR P-PSDB; AAY58634.
XX PR New human polypeptides that regulate gene expression, for treatment,
XX PR prevention and diagnosis of, e.g. cancer.
XX PS Claim 9; Page 144-145; 150pp; English.
XX CC The present sequence is that of Incyte clone 1399169 encoding new human
XX CC protein regulating gene expression PRGE-27 (see AAY58634). The cDNA was
XX CC initially isolated from brain tumour cDNA library BRAITUT08, and the full
XX CC length sequence assembled from overlapping clones from a number of
XX CC libraries. PRGE-27 is expressed in haematopoietic/immune, reproductive
XX CC and gastrointestinal tissues associated with cell proliferative and
XX CC inflammation diseases, disorders or conditions. It is characterized as a
XX CC bromodomain protein. The invention provides PRGE polypeptides (see
XX CC AAY58608-38) and polynucleotides (see AAZ57839-69), expression vectors,
XX CC host cells, antibodies, agonists and antagonists. It also provides
XX CC methods for diagnosing, treating or preventing disorders associated with
XX CC expression of PRGE. Polynucleotides are also used as sources of probes
XX CC and primers for diagnosis and monitoring of disease, also for detecting
XX CC related sequences and in gene mapping
XX SQ Sequence 3190 BP; 918 A; 739 C; 822 G; 711 T; 0 U; 0 Other;
Query Match 74.4%; Score 18.6; DB 3; Length 3190;
Best Local Similarity 84.0%; Pred. No. 3.2e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 AAAAAACACACCACCTGAGCTGGG 25
Db 3138 AAAAAACACACCACCTGAGTGAATGTG 3162
RESULT 37
ADX05939
ID ADX05939 standard; DNA; 3226 BP.
XX AC ADX05939;
XX DT 21-APR-2005 (first entry)
XX DE Cyclin-dependent kinase modulation biomarker DNA SEQ ID NO 504.
XX KW cytosstatic; cyclin-dependent kinase; cdk; biomarker; gene; ds.
XX OS Homo sapiens.
XX PN WO2005012875-A2.
XX PR 10-FEB-2005.
XX PD
XX PF 29-JUL-2004; 2004WO-US024424.
XX PR 29-JUL-2003; 2003US-0490890P.
XX PA (BRIM ) BRISTOL-MYERS SQUIBB CO.
XX PI Li M, Rupnow BA, Webster KR, Jackson DG, Wong TW;
```

```
XX WPI; 2005-163068/17.
XX PR P-PSDB; ADX05940.
XX PT Biomarkers useful for predicting or determining the response of a mammal
XX PT to a cancer treatment comprising administration of a modulator of cyclin-
XX PT dependent kinase activity.
XX PS Claim 5; SEQ ID NO 504; 141pp; English.
XX CC This invention describes a novel method of predicting or determining
XX CC whether a mammal will respond or is responding to an anti-cancer agent
XX CC that modulates cyclin-dependent kinase (cdk) activity. The method
XX CC comprises measuring the level of one or more biomarkers selected from
XX CC 2774 biomarkers given in the specification (nucleotide sequence SEQ ID
XX CC NO:1246 (Genbank EST W28729) is especially preferred). The method of the
XX CC invention is utilized in a kit for determining or predicting whether
XX CC patient would be susceptible or resistant to treatment by an agent
XX CC modulating cdk activity. The invention also describes a method for
XX CC utilizing individualized genetic profiles for treating diseases and
XX CC disorders based on patient's response and molecular level, specialized
XX CC microarrays comprising the biomarkers described, antibodies directed
XX CC against the biomarkers and a cell culture model to identify biomarkers.
XX CC The cdk modulator is preferably N-5-[(5-(1,1-Dimethylethyl)-2-
XX CC oxazolyl)methyl]thio]-2-thiazolyl-4-piperidine carboxamide, 0.5-L-
XX CC tartaric acid salt. Note: The sequence data for this patent did not form
XX CC part of the printed specification, but was obtained in electronic format
XX CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences. This
XX CC sequence encodes a biomarker used in the method of the invention.
XX SQ Sequence 3226 BP; 940 A; 743 C; 828 G; 715 T; 0 U; 0 Other;
Query Match 74.4%; Score 18.6; DB 14; Length 3226;
Best Local Similarity 84.0%; Pred. No. 3.2e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 AAAAAACACACCACCTGAGCTGGG 25
Db 3153 AAAAAACACACCACCTGAGTGAATGTG 3177
RESULT 38
ACN42607
ID ACN42607 standard; cDNA; 3512 BP.
XX AC ACN42607;
XX DT 18-NOV-2004 (first entry)
XX DE Human diagnostic and therapeutic polynucleotide SEQ ID NO:1482.
XX KW ss; gene; gene therapy; human diagnostic and therapeutic polynucleotide;
XX KW dithp.
XX OS Homo sapiens.
XX PN WO2004023973-A2.
XX PR 25-MAR-2004.
XX PD
XX PF 12-SEP-2003; 2003WO-US028227.
XX PR 12-SEP-2002; 2002US-0410259P.
XX PR 12-SEP-2002; 2002US-0410260P.
XX PA (INCY-) INCYTE CORP.
XX PI Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F;
XX PI Harthshorne TA, Suchorolski MT, Altus CM, Pitts SJ, Elder LV;
XX PI Mooney EM, Delegeane AM, Panesar IS, Banville SC, Reddy TP;
XX PI Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstin EH;
XX PI Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LJ;
XX PI Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirton ES;
```

PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.

one of the polynucleotides cited above. The polypeptides and polynucleotides of the invention have antiinflammatory, cytostatic, and antimicrobial activities. The novel polynucleotide may be used to treat disorders by gene therapy. The polypeptides and polynucleotides are useful for treating inflammation, leukaemias, nervous system disorders, or infections. This polynucleotide sequence represents contiguous DNA derived from one of the 235 novel isolated polynucleotides of the invention.

Sequence 16578 BP; 4635 A; 3859 C; 3976 G; 4108 T; 0 U; 0 Other;

Query Match 74.4%; Score 18.6; DB 13; Length 16578;
Best Local Similarity 84.0%; Pred. No. 4.1e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAACAACCACTGAGCTGGG 25
||||| ||||||| |||
Db 16262 AAAAAACAACCACTGCGGTG 16238

RESULT 43
ABZ74034/C
ID ABZ74034 standard; DNA; 39119 BP.
XX AC
XX AC ABZ74034;
XX DT 12-MAY-2003 (first entry)
XX XX
XX DE Secreted protein gene 163 genomic fragment HJPC08, SEQ ID NO:1181.
XX XX
XX KW Human; secreted protein; cancer; tumour; hyperproliferative disorder;
KW autoimmune disorder; inflammation; angiogenic diseases; AIDS;
KW acquired immunodeficiency syndrome; hepatitis; anaemia; wound healing;
KW drug screening; chromosome identification; chromosome mapping;
KW cytostatic; gene therapy; antiinflammatory; immunomodulator; anti-HIV;
KW antianaemic; vulnery; gene; ds.
XX XX
XX OS Homo sapiens.
XX XX
XX PN W0200277013-A2.
XX XX
XX PD 03-OCT-2002.
XX XX
XX PF 26-MAR-2002; 2002WO-US009370.
XX XX
XX PR 27-MAR-2001; 2001US-0278650P.
XX PR 12-SEP-2001; 2001US-00950082.
XX PR 12-SEP-2001; 2001US-00950083.
XX XX
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX XX
XX PI Rosen CA, Ruben SM;
XX DR WPI; 2003-040578/03.
XX XX
XX PT New human secreted proteins and nucleic acids, useful for detecting or
PT treating cancer or other hyperproliferative disorders, autoimmune
PT disorders, inflammatory disorders, HIV disease, hepatitis or anemia.
XX XX
XX PS Disclosure; Page 1836-1845; 2474pp; English.
XX XX
XX CC ABZ73281-ABZ73697 represent cDNAs corresponding to 391 human secreted
CC protein genes, and ABP00947-ABP01363 represent the proteins they encode.
CC CC ABZ73698-ABZ74687 represent human secreted protein genomic fragments. The
CC invention also encompasses antibodies specific for the secreted proteins,
CC the use of the secreted proteins in drug screening and recombinant
CC vectors and host cells comprising a nucleic acid of the invention. The
CC secreted proteins are thought to be involved in biological activities
CC associated with cellular signalling, cellular differentiation, cell
CC migration, prohormone activation and neurotransmitter activity. The
CC secreted proteins, nucleic acids encoding them, antibodies or antibody
CC fragments specific for the secreted proteins, and modulators of protein
CC activity are useful for diagnosing or treating cancers or other

CC	hyperproliferative disorders. Additionally, the secreted proteins and
CC	their nucleic acids may also be used in the treatment of autoimmune
CC	disorders, inflammatory disorders, diseases involving angiogenesis, AIDS
CC	(acquired immunodeficiency syndrome), hepatitis, anaemia, and to promote
CC	wound healing. Nucleic acids of the invention may be used for chromosome
CC	identification, chromosome mapping, in gene therapy, for identifying
CC	individuals from minute biological samples, as hybridisation probes, and
CC	as molecular weight markers. The present sequence represents a human
CC	secreted protein genomic fragment referred to in the disclosure of the
CC	invention
XX	
SQ	Sequence 39119 BP; 11590 A; 8925 C; 7906 G; 10598 T; 0 U; 0 Other;
	Query Match 74.4%; Score 18.6; DB 8; Length 39119;
	Best Local Similarity 84.0%; Pred. No. 4.7e+02;
	Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Qy	1 AAAAAACACCCACTGAGCTGGG 25
Db	17118 AAAAAACACCCAGTGAATGTG 17094
RESULT 44	
ID	ADA98641/c
ID	ADA98641 standard; DNA; 39119 BP.
XX	
AC	ADA98641;
XX	
DT	20-NOV-2003 (first entry)
XX	
DE	Human secreted protein-related DNA sequence #234.
XX	
KW	human; secreted protein; cardiovascular disorder; arrhythmia;
KW	atherosclerosis; stroke; endocarditis; congestive heart failure;
KW	rheumatic heart disease; cardiomyopathy; hemorrhoids; varicose veins;
KW	migraine; thrombosis; neural disorder; immune system disorder;
KW	muscular disorder; reproductive disorder; gastrointestinal disorder;
KW	pulmonary disorder; renal disorder; proliferative disorder; cancer; ds.
XX	
OS	Homo sapiens.
XX	
FN	WO2003004623-A2.
XX	
PD	16-JAN-2003.
XX	
PF	26-MAR-2002; 2002WO-US009922.
XX	
PR	27-MAR-2001; 2001US-0278650P.
PR	12-SEP-2001; 2001US-00950082.
PR	12-SEP-2001; 2001US-00950083.
XX	
PA	(HUMA-) HUMAN GENOME SCI INC.
XX	
FI	Rosen CA, Ruben SM;
XX	
DR	WPI; 2003-247946/24.
XX	
PT	New human secreted polypeptide and nucleic acid molecules, useful for
PT	diagnosing, preventing, prognosticating or treating cardiovascular
PT	disorders (e.g. arrhythmia, atherosclerosis, cardiomyopathy, or
PT	thrombosis).
XX	
PS	Disclosure; SEQ ID NO 750; 1572pp; English.
XX	
CC	The invention comprises the amino acid and coding sequence of human
CC	secreted proteins. The DNA and protein sequences of the invention are
CC	useful in the treatment of cardiovascular disorders, such as: arrhythmia,
CC	atherosclerosis, stroke, endocarditis, congestive heart failure,
CC	rheumatic heart disease, cardiomyopathy, hemorrhoids, varicose veins,
CC	migraine, or thrombosis. The DNA and protein sequences may also be used
CC	for treating or preventing: neural disorders, immune system disorders,
CC	muscular disorders, reproductive disorders, gastrointestinal disorders,
CC	pulmonary disorders, renal disorders, proliferative disorders and/or

```
CC cancerous diseases. The present DNA sequence is used in the
CC exemplification of the invention. NOTE: The present sequence is shown on
CC the WIPO website.
XX
SQ Sequence 39119 BP; 11690 A; 8925 C; 7906 G; 10598 T; 0 U; 0 Other;
Query Match 74.4%; Score 18.6; DB 8; Length 39119;
Best Local Similarity 84.0%; Pred. No. 4.7e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
OY 1 AAAAAACACCACTGAGCTGGG 25
Db 17118 AAAAAACACCACTGAGCTGGG 17094
RESULT 45
ID ADC20764/c
XX ID ADC20764 standard; DNA; 39119 BP.
XX AC ADC20764;
XX DT 18-DEC-2003 (first entry)
XX DE Human secreted protein-related DNA sequence #182.
XX KW gene therapy; human; secreted protein; haemopoietic disorder;
KW haematological disorder; anaemia; haemophilia; inflammatory disorder;
KW inflammatory bowel disease; Crohn's disease; neoplastic disease; cancer;
KW leukaemia; wound healing; epithelial cell proliferation disorder;
KW immune disorder; autoimmune disorder; asthmatic disorder;
KW cardiovascular disorder; atherosclerosis; myocarditis;
KW infectious disease; HIV; AIDS; endocrine disorder; diabetes;
KW gastrointestinal disorder; duodenal ulcer; gastroenteritis; gene; ds.
XX OS Homo sapiens.
XX PN WO200292787-A2.
XX PD 21-NOV-2002.
XX PF 26-MAR-2002; 2002WO-US009257.
XX PR 27-MAR-2001; 2001US-0278650P.
XX PR 12-SEP-2001; 2001US-00950082.
XX PR 12-SEP-2001; 2001US-00950082.
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PI Rosen CA, Ruben SM;
XX WPI; 2003-129287/12.
XX
XX New human secreted proteins and nucleic acid molecules, useful for
XX preparing a diagnostic or pharmaceutical composition for diagnosing,
XX preventing or treating hematopoietic or hematologic disorders, e.g.
XX anemia or hemophilia.
XX
XX Disclosure; SEQ ID NO 718; 1512pp; English.
XX
XX The invention comprises the amino acid and coding sequences of human
XX secreted proteins. The DNA and protein sequences of the invention are
XX useful for detecting, preventing, diagnosing, prognosticating, treating
XX or ameliorating; haematopoietic or haematological disorders (e.g. anaemia
XX and hemophilia); inflammatory disorders (e.g. inflammatory bowel disease
XX and Crohn's disease); neoplastic disease (e.g. cancer and leukaemia);
XX wound healing and disorders of epithelial cell proliferation; immune
XX disorders (e.g. autoimmune disorders and asthmatic disorders);
XX cardiovascular disorders (e.g. atherosclerosis and myocarditis);
XX infectious disease (e.g. HIV/AIDS); endocrine disorders (e.g. diabetes);
XX and gastrointestinal disorders (e.g. duodenal ulcers and
XX gastroenteritis). The present DNA sequence was used in the
XX exemplification of the invention.
XX
SQ Sequence 39119 BP; 11690 A; 8925 C; 7906 G; 10598 T; 0 U; 0 Other;
Query Match 74.4%; Score 18.6; DB 10; Length 39119;
Best Local Similarity 84.0%; Pred. No. 4.7e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
OY 1 AAAAAACACCACTGAGCTGGG 25
Db 17118 AAAAAACACCACTGAGCTGGG 17094
RESULT 46
ID ABZ67621/c
XX ID ABZ67621 standard; DNA; 39119 BP.
XX AC ABZ67621;
XX DT 26-MAR-2003 (first entry)
XX DE Human secreted protein encoding genomic DNA SEQ ID NO 1144.
XX KW Human; secreted protein; neutropenic; neuroprotective; cytostatic;
KW viricide; dermatological; immunosuppressive; antiinflammatory; anti-HIV;
KW vulnery; antibacterial; antiparkinsonian; antisking; antianemic;
KW antiarthritic; cancer; antirheumatic; hepatotropic; cerebroprotective;
KW antiinflammatory; antiallergic; antidiabetic; antitumor; anticonvulsant;
KW antifungal; antiparasitic; cardiant; immune disorder; infection; vaccine;
KW cardiovascular disorder; neurological disease; nephrotropic;
KW gene therapy; gene; ds.
XX OS Homo sapiens.
XX PN WO200277186-A2.
XX PD 03-OCT-2002.
XX PF 26-MAR-2002; 2002WO-US009188.
XX PR 27-MAR-2001; 2001US-0278650P.
XX PR 12-SEP-2001; 2001US-00950082.
XX PR 12-SEP-2001; 2001US-00950083.
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PI Rosen CA, Ruben SM;
XX WPI; 2003-040583/03.
XX
XX New human secreted proteins encoded by genes contained in cDNA clones
XX (e.g. HGAC19), useful for preventing, treating or diagnosing e.g. AIDS,
XX multiple sclerosis, herpes virus, leukemia, tick-borne encephalitis or
XX West Nile fever.
XX
XX Disclosure; Page 1804-1814; 2423pp; English.
XX
XX The invention relates to novel human genes (ABZ66891-ABZ68209) and the
XX encoded secreted proteins (ABP99470-ABP99872) useful for preventing,
XX treating or ameliorating medical conditions e.g. by protein or gene
XX therapy. The genes are isolated from a range of human tissues disclosed
XX in the specification. The nucleic acids, proteins, antibodies and
XX (ant)agonists are useful in the diagnosis, treatment and prevention of:
XX (a) cancer, e.g. breast and ovarian cancer and other cancers of the
XX adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver,
XX lung or urogenital; (b) immune disorders e.g. Addison's disease,
XX allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
XX diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
XX arthritis and ulcerative colitis; (c) cardiovascular disorders such as
XX myocardial ischaemias; (d) wound healing; (e) neurological diseases e.g.
XX cerebral anoxia and epilepsy; and (f) infectious diseases such as viral,
XX bacterial, fungal and parasitic infections
XX
XX Sequence 39119 BP; 11690 A; 8925 C; 7906 G; 10598 T; 0 U; 0 Other;
```

```

Query Match      74.4%; Score 18.6; DB 10; Length 39119;
Best Local Similarity 84.0%; Pred. No. 4.7e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 AAAAAACAACACACTGAGCTGGG 25
Db 17118 AAAAAACAACACACTGAGCTGG 17094

RESULT 47
ABS98185/c
ID ABS98185 standard; DNA; 53099 BP.
XX
AC ABS98185;
XX
DT 23-DEC-2002 (first entry)
XX
DE Human multidrug resistance gene exons 4-28 sequence.
XX
KW Human; db; cytochrome P450 A1; CYP450A1; UGT2B4; MDR1;
KW cytochrome P450 A2; CYP450A2; cytochrome P450 02E; CYP45002E1; LTF;
KW adrenergic receptor beta1; ADRB1; aryl hydrocarbon; AHR; MRP3; NR1I2;
KW aryl hydrocarbon receptor nuclear translocator; ARNT; cathepsin S; CTSS;
KW cyclooxygenase 2; COX2; diazepam binding inhibitor; DBI; haematological;
KW epoxide hydroxylase 2; EPHX2; 5-lipoxygenase activating protein; FLAP;
KW glutathione-S-transferase 12; GST12; histamine-N-methyl transferase;
KW HMT; kallikrein 2; KLK2; nicotinamide-N-methyl transferase; NNMT;
KW NADPH quinone oxidoreductase 2; NQO2; sulfoltransferase thermolabile; STM;
KW UDP-glucuronosyl transferase 2B4; UDP-glucuronosyl transferase 2B7;
KW UGT2B7; UDP-glucuronosyl transferase; UGT2B15; urokinase receptor; uPA;
KW multidrug resistance 1; lactotransferrin; orphan nuclear receptor;
KW multidrug resistance associated protein 3; cancer; prostate;
KW acetylcholine muscarinic receptor; CHMR1; CHMR2; CHMR3; CHMR4; CHMR5;
KW altered drug metabolism; cardiovascular function; colorectal tumour;
KW central nervous system; pulmonary; immunological; SNP;
KW single nucleotide polymorphism.
XX
OS Homo sapiens.
XX
FN WO200257410-A2.
XX
PD 25-JUL-2002.
XX
PF 28-NOV-2001; 2001WO-US044838.
XX
PR 28-NOV-2000; 2000US-00724389.
XX
PA (DNAS-) DNA SCI LAB INC.
XX
PI Guida M, Hall J;
XX
DR WPI; 2002-698522/75.
XX
PT Isolated nucleic acid molecules having polymorphisms in known human genes
PT e.g. cytochrome P450 and cathepsin S useful as genetic linkage markers
PT for locating, identifying and characterizing the genes responsible for
PT disorder-related traits.
XX
PS Example 22; Page 437-467; 714pp; English.
XX
CC This invention relates to the sequence of an isolated nucleic acid
CC molecule comprising at least one base variation from that of a known
CC human cytochrome P450 A1 (CYP450A1), cytochrome P450 A2 (CYP450A2),
CC cytochrome P450 02E1 (CYP45002E1), adrenergic receptor beta1 (ADRB1),
CC aryl hydrocarbon (AHR), aryl hydrocarbon receptor nuclear translocator
CC (ARNT), cathepsin S (CTSS), cyclooxygenase 2 (COX2), diazepam binding
CC inhibitor (DBI), epoxide hydroxylase 2 (EPHX2), 5-lipoxygenase activating
CC protein (FLAP), glutathione-S-transferase 12 (GST12), histamine-N-methyl
CC transferase (NNMT), (kallikrein 2) KLK2, nicotinamide -N-methyl
CC sulfoltransferase thermolabile (STM), UDP-glucuronosyl transferase 2B4
CC (UGT2B4), UDP-glucuronosyl transferase 2B7 (UGT2B7), UDP-glucuronosyl
CC transferase (UGT2B15), urokinase receptor (uPA), multidrug resistance 1

```

```

(MDR1), lactotransferrin (LTF), multidrug resistance associated protein 3
(MRP3), orphan nuclear receptor (NR1I2), or acetylcholine muscarinic
receptor 1, 2, 3, 4, or 5 (CHMR1, CHMR2, CHMR3, CHMR4 or CHMR5) sequence.
The polymorphisms in the human genes cited in the invention are useful as
genetic linkage markers for locating and characterizing the genes that
are responsible for specific traits within the genome and eventually
identifying the genes responsible for a variety of disorder-related
traits as a result of their e.g., overexpression, constitutive
expression, mutation or underexpression, which may be used in diagnosing
and/or treating the disorders. The nucleic acid molecules comprising the
polymorphic sequences contained in CYP450A1, CYP450A2, CYP4502E1,
ARNT, EPHX2, GST12, NNMT, NQO2, NR1I2, STM, UGT2B4, UGT2B7, UGT2B15, AHR,
MDR1 and/or MDR3 are useful for screening individuals for altered drug
metabolism. The polymorphic sequences contained in CYP450A1, CYP450A2,
AHR, MDR1 and/or MDR3 may also be used to screen individuals for
susceptibility to cancer. Polymorphic sequences in ADRB1 or CHMR2 are
used to screen for altered cardiovascular function, in COX2 for altered
susceptibility to colorectal tumours, in DBI or CHMR1 for altered central
nervous system function, in FLAP and NNMT for altered pulmonary,
immunological or haematological function, in KLK2 for altered serine
protease activity in the prostate, in LTF for altered immunological or
haematological function, in CHMR3, CHMR4 or CHMR5 for altered central and
peripheral nervous system function. The present sequence represents a
polymorphic DNA sequence of the invention
XX
SQ Sequence 53099 BP; 16225 A; 9740 C; 10096 G; 17038 T; 0 U; 0 Other;

Query Match      74.4%; Score 18.6; DB 6; Length 53099;
Best Local Similarity 84.0%; Pred. No. 4.9e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 AAAAAACAACACACTGAGCTGGG 25
Db 7139 AAAAAACAACACTGAGCTGGG 7115

RESULT 48
ADX98572
ID ADX98572 standard; DNA; 86950 BP.
XX
AC ADX98572;
XX
DT 05-MAY-2005 (first entry)
XX
DE Human HT014/LOC148902/LYPLA2/GALE genomic DNA.
XX
KW SNP detection; breast tumor; endocrine disease;
KW gynecology and obstetrics; neoplasm; cytostatic; metastasis;
KW gene therapy; RNA interference; chromosome 1; ds; SNP;
KW single nucleotide polymorphism; UDP-galactose 4-epimerase; GALE;
KW UDP-glucose 4-epimerase; lysophospholipase; HT014; LOC148902; LYPLA2.
XX
OS Homo sapiens.
XX
FH Key
FH variation 215
FH /tag= a
FH /standard_name= "Single nucleotide polymorphism (SNP)"
FH variation 385
FH /tag= b
FH /standard_name= "Single nucleotide polymorphism (SNP)"
FH variation 2019
FH /tag= c
FH /standard_name= "Single nucleotide polymorphism (SNP)"
FH variation 3112
FH /tag= d
FH /standard_name= "Single nucleotide polymorphism (SNP)"
FH variation 3437
FH /tag= e
FH /standard_name= "Single nucleotide polymorphism (SNP)"
FH variation 4326
FH /tag= f
FH /standard_name= "Single nucleotide polymorphism (SNP)"

```

FT	variation	4523	/*tag= g	/*standard_name= "Single nucleotide polymorphism (SNP) "	FT	/*tag= ae	/*standard_name= "Single nucleotide polymorphism (SNP) "
FT	variation	4645	/*tag= h	/*standard_name= "Single nucleotide polymorphism (SNP) "	FT	18538	/*tag= af
FT	variation	4646	/*tag= i	/*standard_name= "Single nucleotide polymorphism (SNP) "	FT	18810	/*tag= ag
FT	variation	5830	/*tag= j	/*standard_name= "Single nucleotide polymorphism (SNP) "	FT	20282	/*tag= ah
FT	variation	6556	/*tag= k	/*standard_name= "Single nucleotide polymorphism (SNP) "	FT	20495	/*tag= ai
FT	variation	6568	/*tag= l	/*standard_name= "Single nucleotide polymorphism (SNP) "	FT	20633	/*tag= aj
FT	variation	6570	/*tag= m	/*standard_name= "Single nucleotide polymorphism (SNP) "	FT	20659	/*tag= ak
FT	variation	9351	/*tag= n	/*standard_name= "Single nucleotide polymorphism (SNP) "	FT	21603	/*tag= al
FT	variation	10874	/*tag= o	/*standard_name= "Single nucleotide polymorphism (SNP) "	FT	24075	/*tag= am
FT	variation	10926	/*tag= p	/*standard_name= "Single nucleotide polymorphism (SNP) "	FT	24191	/*tag= an
FT	variation	11462	/*tag= q	/*standard_name= "Single nucleotide polymorphism (SNP) "	FT	25010	/*tag= ao
FT	variation	12415	/*tag= r	/*standard_name= "Single nucleotide polymorphism (SNP) "	FT	25124	/*tag= ap
FT	variation	12668	/*tag= s	/*standard_name= "Single nucleotide polymorphism (SNP) "	FT	25145	/*tag= aq
FT	variation	13886	/*tag= t	/*standard_name= "Single nucleotide polymorphism (SNP) "	FT	25222	/*tag= ar
FT	variation	15005	/*tag= u	/*standard_name= "Single nucleotide polymorphism (SNP) "	FT	25288	/*tag= as
FT	variation	15387	/*tag= v	/*standard_name= "Single nucleotide polymorphism (SNP) "	FT	26080	/*tag= at
FT	variation	15639	/*tag= w	/*standard_name= "Single nucleotide polymorphism (SNP) "	FT	26556	/*tag= au
FT	variation	15761	/*tag= x	/*standard_name= "Single nucleotide polymorphism (SNP) "	FT	27416	/*tag= av
FT	variation	16969	/*tag= y	/*standard_name= "Single nucleotide polymorphism (SNP) "	FT	27432	/*tag= aw
FT	variation	17124	/*tag= z	/*standard_name= "Single nucleotide polymorphism (SNP) "	FT	29841	/*tag= ax
FT	variation	17369	/*tag= aa	/*standard_name= "Single nucleotide polymorphism (SNP) "	FT	29955	/*tag= ay
FT	variation	17523	/*tag= ab	/*standard_name= "Single nucleotide polymorphism (SNP) "	FT	30160	/*tag= az
FT	variation	18158	/*tag= ac	/*standard_name= "Single nucleotide polymorphism (SNP) "	FT	30294	/*tag= ba
FT	variation	18220	/*tag= ad	/*standard_name= "Single nucleotide polymorphism (SNP) "	FT	31039	/*tag= bb
FT	variation	18224	/*tag= ae	/*standard_name= "Single nucleotide polymorphism (SNP) "	FT	32071	/*tag= bc

FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT 33698 /tag= bd
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT 35555 /tag= be
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT 36658 /tag= bf
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT 37720 /tag= bg
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT 37937 /tag= bh
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT 38304 /tag= bi
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT 42395 1 AAAAAACACCACCACTGAGCTGGG 25
DB 42395 AAAAAAGATTACCACTGGGCTGGG 42419

Query Match 74.4%; Score 18.6; DB 14; Length 86950;
Best Local Similarity 84.0%; Pred. No. 5.3e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

RESULT 49
ACN44026/C
ID ACN44026 standard; DNA; 122923 BP.
XX ACN44026;
XX 18-NOV-2004 (first entry)
XX Human genomic sequence hCG24510.
XX Cytostatic; carcinoma; lymphoma; cancer; human; gene; ss.
XX Homo sapiens.
XX WO2003073826-A2.
XX 12-SEP-2003.
XX 28-FEB-2003; 2003WO-US006235.
XX 01-MAR-2002; 2002US-00087192.
XX (SAGR-) SAGRES DISCOVERY.
XX Morris DW;
XX WPI; 2003-328604/31.
XX Recombinant nucleic acid useful for diagnosis and treatment of carcinoma comprises a nucleotide sequence.
XX Claim 1; SEQ ID NO 268; Opp; English.
XX The present invention relates to novel DNA and protein sequences which are associated with carcinomas. The sequences are useful for: (i) for screening drug candidates; (ii) for screening of bioactive agent capable of binding to Carcinoma Associated Protein (CAP); (iii) for screening of a bioactive agent capable of modulating the activity of CAP; (iv) for evaluating the effect of a candidate carcinoma drug; (v) for diagnosing carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip; (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for determining Carcinoma Associated (CA) gene copy number. In addition, the CA genes are useful as DNA vaccines and the CAP are useful as markers of

CC carcinoma including lymphoma. The present sequence is one such CA coding sequence. Note: This patent is an equivalent to basic patent
CC US2002182586A1, for which no sequence data was published
XX SQ Sequence 122923 BP; 28628 A; 24391 C; 25218 G; 32559 T; 0 U; 12127 Other;
Query Match 74.4%; Score 18.6; DB 11; Length 122923;
Best Local Similarity 84.0%; Pred. No. 5.6e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAACACCACCACTGAGCTGGG 25
DB 12267 AAAAAAACACACAGAGCTAGG 12243

RESULT 50
ACF62749
ID ACF62749 standard; DNA; 128993 BP.
XX ACF62749;
XX 08-OCT-2003 (first entry)
XX Cancer based on CYP3A5 related polynucleotide SEQ ID NO:681.
XX Cancer; CYP3A5; irinotecan; pharmaceutical; malignant glioma;
KW cytochrome p450; subfamily IIIA; nifedipine oxidase; polypeptide 5;
KW cytostatic; gene; ds.
XX Unidentified.
XX WO2003013534-A2.
XX 20-FEB-2003.
XX 23-JUL-2002; 2002WO-EP008219.
XX 23-JUL-2001; 2001EP-00117608.
PR 24-MAY-2002; 2002EP-00011710.
XX (EPID-) EPIDAUROS BIOTECHNOLOGIE AG.
XX Heinrich G, Kerb R;
XX WPI; 2003-268144/26.
XX New use of irinotecan for preparation of compositions for treating cancer in subject having genome with variant allele comprising cytochrome p450, subfamily IIIA, polypeptide 5 polynucleotide, termed CYP3A5.
XX Disclosure; SEQ ID NO 681; 86pp; English.
XX The present invention describes the use of irinotecan (I) or its derivative for the preparation of a pharmaceutical composition for treating colorectal, cervical, gastric, lung, ovarian or pancreatic cancer, or malignant glioma in a subject having a genome with a variant allele which comprises a cytochrome p450, subfamily IIIA (nifedipine oxidase), polypeptide 5 (CYP3A5) polynucleotide (II). (I) and (II) have cytostatic activity. The therapeutic applications of (I) is improved, since it is possible to individually treat a subject with an appropriate dosage and/or an appropriate derivative of (I). Therefore, undesirable, harmful or toxic effects are efficiently avoided. Unnecessary and potentially harmful treatment of those subjects who do not respond to the treatment with substances (nonresponders), as well as the development of drug resistances due to suboptimal drug dosing can be avoided. ACF62200 to ACF62751 and ABM34912 to ABM35013 represent sequences used in the exemplification of the present invention
XX SQ Sequence 128993 BP; 41329 A; 25384 C; 24283 G; 37983 T; 0 U; 14 Other;
Query Match 74.4%; Score 18.6; DB 8; Length 128993;
Best Local Similarity 84.0%; Pred. No. 5.6e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 AAAAACAACCACTGAGCTGG 25
Db 121849 AAAAGACACCCCTCACTGAGCTGG 121873

Search completed: February 3, 2006, 21:56:24
Job time : 212.111 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 3, 2006, 14:38:25 ; Search time 2951.67 Seconds
(without alignments)
396.277 Million cell updates/sec

Title: US-10-719-900-3

Perfect score: 25

Sequence: 1 aaaaacaccaccactgagctggg 25

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database :

EST:*

- 1: gb_est1:*
- 2: gb_est2:*
- 3: gb_est3:*
- 4: gb_est4:*
- 5: gb_est5:*
- 6: gb_est6:*
- 7: gb_est7:*
- 8: gb_est8:*
- 9: gb_est9:*
- 10: gb_est10:*
- 11: gb_est11:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	25	100.0	403	1	AI551075
2	25	100.0	412	5	BY627047
3	25	100.0	444	1	AW909560
4	25	100.0	499	6	CB249608
5	25	100.0	517	3	BM935380
6	25	100.0	664	1	BB035891
7	25	100.0	1993	4	AK077834
8	25	100.0	3000	4	AK086675
9	23.4	93.6	430	5	BY635162
10	23.4	93.6	450	2	BF464082
11	23.4	93.6	478	2	BB760039
12	22	88.0	395	5	BY650318
13	21	84.0	394	5	BY674654
14	20.2	80.8	172	7	CV525379
15	20.2	80.8	256	7	CV525125
16	20.2	80.8	375	5	BY64584
17	20.2	80.8	577	1	AL633505
18	20.2	80.8	596	5	EX746867
19	20.2	80.8	600	5	BQ526123
20	20.2	80.8	606	1	AL681607
21	20.2	80.8	653	1	AL860221
22	20.2	80.8	654	1	AL854589

```
c 96 18.8 75.2 485 9 A0845099 an46a08 J
c 97 18.8 75.2 486 9 C983799 ZUAC988TH
c 98 18.8 75.2 495 6 CB713124 AMGNNUC.N
c 99 18.8 75.2 501 9 A0845100 an46a09 J
c 100 18.8 75.2 510 9 A0844270 an29c08 J
c 101 18.8 75.2 519 9 A0845079 an44g06 J
c 102 18.8 75.2 530 6 CB719046 AMGNNUC.N
c 103 18.8 75.2 543 10 C2341486 ZMWBFO111
c 104 18.8 75.2 559 9 BZ776093 ih98d01.g
c 105 18.8 75.2 577 9 BH182062 020_O_19-
c 106 18.8 75.2 577 11 CNS07NJM AL619012 T3 end of
c 107 18.8 75.2 581 10 CG024655 ZMWBBC056
c 108 18.8 75.2 586 6 CB077287 hj52a03.g
c 109 18.8 75.2 624 10 C0793416 ZMWBBD030
c 110 18.8 75.2 632 10 CW013690 ZMWBLC000
c 111 18.8 75.2 645 6 CA351434 622504 NC
c 112 18.8 75.2 645 8 CG029229 1341104 N
c 113 18.8 75.2 653 9 A2416349 IM0191K05
c 114 18.8 75.2 656 10 CL355109 RPCI44_40
c 115 18.8 75.2 676 10 C0736482 ZMWBBD032
c 116 18.8 75.2 692 10 AG126291 Pan trogl
c 117 18.8 75.2 698 10 CG010669 ZUAEV47TV
c 118 18.8 75.2 713 9 CC977408 ZUAE725TV
c 119 18.8 75.2 713 11 CR867166 Sus scrof
c 120 18.8 75.2 715 6 CA368573 644796 NC
c 121 18.8 75.2 725 9 CC461247 ZMWBBC036
c 122 18.8 75.2 727 9 BZ651474 OGBAD90TM
c 123 18.8 75.2 733 9 BZ988638 PUCAW06TD
c 124 18.8 75.2 737 10 CG795584 ZMWBBD032
c 125 18.8 75.2 745 10 CW004239 ZMWBLa000
c 126 18.8 75.2 752 6 CD102779 AGENCOURT
c 127 18.8 75.2 766 10 C2274186 ZMWBFO055
c 128 18.8 75.2 770 10 CG872643 ZMWBBC027
c 129 18.8 75.2 775 9 CC175852 ZMWBBC029
c 130 18.8 75.2 782 11 CR867346 Sus scrof
c 131 18.8 75.2 792 10 C2350903 ZMWBFO086
c 132 18.8 75.2 798 9 CC753941 ZMWBBD013
c 133 18.8 75.2 807 10 C2304352 ZMWBFO082
c 134 18.8 75.2 809 7 CK316740 SB02019A2
c 135 18.8 75.2 813 9 CC703882 OGUET43TH
c 136 18.8 75.2 819 10 C2388932 ZMWBFO165
c 137 18.8 75.2 828 10 C2338950 ZMWBFO107
c 138 18.8 75.2 837 9 CC333520 OG1BE01TH
c 139 18.8 75.2 845 9 CC974619 ZUAHE63TV
c 140 18.8 75.2 848 10 C2296777 ZMWBFO066
c 141 18.8 75.2 849 9 CC825858 ZMWBBD017
c 142 18.8 75.2 849 9 CC834256 ZMWBBD018
c 143 18.8 75.2 853 10 C2276273 ZMWBFO059
c 144 18.8 75.2 859 10 CL281555 ZMWBBD062
c 145 18.8 75.2 875 9 CC992763 ZUABE23TV
c 146 18.8 75.2 876 10 CW002921 ZMWBHK000
c 147 18.8 75.2 879 9 CC333526 OG1BE01TV
c 148 18.8 75.2 884 10 CG223383 OG1AD06TH
c 149 18.8 75.2 886 9 CC978249 ZUAHE79TV
c 150 18.8 75.2 886 9 CC978249 ZUAHE79TV
```

ALIGNMENTS

```
RESULT 1
A1551075/c 403 bp mRNA linear EST 23-MAR-1999
LOCUS IMAGE:1277237 3', mRNA sequence.
DEFINITION vx35h03.x1 Stratagene mouse lung 937302 Mus musculus cDNA clone
ACCESSION A1551075
VERSION A1551075.1 GI:4483438
KEYWORDS Mus musculus (house mouse)
SOURCE EST.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Murinae; Mus.
```

```
REFERENCE
AUTHORS
A0845099 an46a08 J
C983799 ZUAC988TH
CB713124 AMGNNUC.N
A0845100 an46a09 J
A0844270 an29c08 J
A0845079 an44g06 J
CB719046 AMGNNUC.N
C2341486 ZMWBFO111
BZ776093 ih98d01.g
BH182062 020_O_19-
AL619012 T3 end of
CG024655 ZMWBBC056
CB077287 hj52a03.g
C0793416 ZMWBBD030
CW013690 ZMWBLC000
CA351434 622504 NC
CG029229 1341104 N
A2416349 IM0191K05
CL355109 RPCI44_40
C0736482 ZMWBBD032
AG126291 Pan trogl
CG010669 ZUAEV47TV
CC977408 ZUAE725TV
CR867166 Sus scrof
CA368573 644796 NC
CL243492 ZMWBBD039
CC461247 ZMWBBC036
BZ651474 OGBAD90TM
BZ988638 PUCAW06TD
CG795584 ZMWBBD032
CW004239 ZMWBLa000
CD102779 AGENCOURT
C2274186 ZMWBFO055
CG872643 ZMWBBC027
CC175852 ZMWBBC029
CR867346 Sus scrof
C2350903 ZMWBFO086
CC753941 ZMWBBD013
C2304352 ZMWBFO082
CK316740 SB02019A2
CC703882 OGUET43TH
C2388932 ZMWBFO165
C2338950 ZMWBFO107
CC333520 OG1BE01TH
CC974619 ZUAHE63TV
C2296777 ZMWBFO066
CC825858 ZMWBBD017
CC834256 ZMWBBD018
C2276273 ZMWBFO059
CL281555 ZMWBBD062
CC992763 ZUABE23TV
CW002921 ZMWBHK000
CC333526 OG1BE01TV
CG223383 OG1AD06TH
CC978249 ZUAHE79TV

RESULT 2
BY627047
LOCUS
DEFINITION cDNA clone K430014A06 3', mRNA sequence.
ACCESSION BY627047
VERSION BY627047.1 GI:26962229
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Murinae; Mus.

REFERENCE
AUTHORS
1 (bases 1 to 412)
Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
Nikaido, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I.,
Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,
Schonbach, C., Gojohori, T., Baldarelli, R., Hill, D.P., Bult, C.,
Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H.,
Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Bruscia, V.,
Chothia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A.,
Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T.,
Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S.,
```

```
1 (bases 1 to 403)
Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T.,
Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y.,
Person, B., Waller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R.,
Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R.,
Waterston, R. and Wilson, R.
The WashU-NCI Mouse EST Project 1999
Unpublished (1999)
Contact: Marra M/WashU-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:669037
This clone was previously sequenced on the 5' end only, this new
data is from the 3' end
Possible reversed clone: polyT not found
High quality sequence stop: 388.
```

FEATURES

```
source
1..403
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6 x CBA"
/db_xref="taxon:10090"
/clone="IMAGE:1277237"
/sex="female"
/tissue_type="lung"
/dev_stage="6-8 month old"
/lab_host="SOLR (kanamycin resistant)"
/clone_lib="Stratagene mouse lung 937302"
/notice="Organ: lung; Vector: pBluescript SK-; Site:1:
EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer:
Oligo dt. 6-8 month old female lung and 1.5 year old male
lung were source of mRNA. Average insert size: 1.5 kb;
Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGACGAG
3' -3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'
```

ORIGIN

```
Query Match 100.0%; Score 25; DB 1; Length 403;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAAAACACCACTGAGCTGGG 25
| | | | | | | | | | | | | | | | | |
Db 292 AAAAAAACACCACTGAGCTGGG 268
```

RESULT 2

```
BY627047
LOCUS
DEFINITION cDNA clone K430014A06 3', mRNA sequence.
```

```
ACCESSION BY627047
VERSION BY627047.1 GI:26962229
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
```

```
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Murinae; Mus.
```

REFERENCE

```
1 (bases 1 to 412)
Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
Nikaido, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I.,
Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,
Schonbach, C., Gojohori, T., Baldarelli, R., Hill, D.P., Bult, C.,
Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H.,
Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Bruscia, V.,
Chothia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A.,
Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T.,
Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S.,
```

Gustincich,S., Hirokawa,N., Jackson,I.J., Jarvis,E.D., Kanai,A., Kawaji,H., Kawasawa,Y., Kedzierski,R.M., King,B.L., Konagaya,A., Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R., Maltais,L., Marchionni,L., McKenzie,L., Miki,H., Nagashima,T., Numata,K., Okido,T., Pavan,W.J., Perteau,G., Pesole,G., Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramachandran,S., Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring,B.Z., Ringwald,M., Sandelin,A., Schneider,C., Semple,C.A., Setou,M., Shimada,K., Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tomita,M., Verardo,R., Wagner,L., Wahlestedt,C., Wang,Y., Watanabe,Y., Wells,C., Wilming,L.G., Wynshaw-Boris,A., Yanagisawa,M., Yang,I., Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A., Carninci,P., Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura,M., Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K., Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata,K., Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S., Rogers,J., Birney,E. and Hayashizaki,Y.

TITLE
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

JOURNAL
Nature 420, 563-573 (2002)

PUBMED
12466851

COMMENT
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Sueniro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216

Email: genome-resesc.riken.jp, URL: <http://genome-gsc.riken.jp/>
Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hirazane,T., Imotani,K., Ishii,Y., Itoh,M., Kawai,J., Konno,H., Miyazaki,A., Murata,M., Nakamura,M., Nomura,K., Numazaki,R., Ohno,M., Sakai,K., Sakazume,N., Sasaki,D., Sato,K., Shibata,K., Shiraki,T., Tagami,M., Waki,K., Watahiki,A., Muramatsu,M. and Hayashizaki,Y. Direct Submission
Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Tissues were provided by Michela Fagiolini and Takao K. Henschi (Laboratory for Neuronal Circuit Development Brain Science Institute RIKEN 2-1 Hiroawa, Wako-shi, Saitama 351-0198 Japan) whose assistance we gratefully acknowledge.
Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES
source
Location/Qualifiers
1..412
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="K430014A06"
/tissue_type="visual cortex"
/clone_lib="RIKEN full-length enriched, visual cortex"

ORIGIN
Query Match 100.0%; Score 25; DB 5; Length 412;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY
Db

RESULT 3
AW909560/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

FEATURES
source
Location/Qualifiers
1..444
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:3156245"
/sex="female, virgin"
/tissue_type="infiltrating ductal carcinoma"
/dev_stage="5 months"
/clone_lib="NCI CGAP Mam6"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: Sali; Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt. Library constructed by Life Technologies. Investigator providing samples: Jeffrey Green, M.D., NIH"

ORIGIN
Query Match 100.0%; Score 25; DB 1; Length 444;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY
Db
RESULT 4
CB249608
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS

1 AAAAAACAACCACTGAGCTGGG 25
|||||
14 AAAAAACAACCACTGAGCTGGG 38
|||||

AW909560 444 bp mRNA linear EST 25-MAY-2000
ur76h03.x1 NCI_CGAP_Mam6 Mus musculus cDNA clone IMAGE:3156245 3',
mRNA sequence.
AW909560
AW909560.1 GI:8074797
EST.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Murinae; Mus.
1 (bases 1 to 444)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Other ESTs: ur76h03.y1
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Jeffrey Green M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
image.llnl.gov/image/html/iresources.shtml

MGI:1059001
Possible reversed clone: polyT not found
Seq primer: -40UP from Gibco
High quality sequence stop: 407.
Location/Qualifiers
1..444
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:3156245"
/sex="female, virgin"
/tissue_type="infiltrating ductal carcinoma"
/dev_stage="5 months"
/clone_lib="NCI CGAP Mam6"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: Sali; Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt. Library constructed by Life Technologies. Investigator providing samples: Jeffrey Green, M.D., NIH"

ORIGIN
Query Match 100.0%; Score 25; DB 1; Length 444;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 AAAAAACAACCACTGAGCTGGG 25
|||||
223 AAAAAACAACCACTGAGCTGGG 199
|||||

CB249608 499 bp mRNA linear EST 15-JUL-2003
UI-M-EXO-by1-g-22-0-UI.r1 NIH_BMAP_EXO Mus musculus cDNA clone
IMAGE:5719725 5', mRNA sequence.
CB249608
CB249608.1 GI:28389208
EST.

```

SOURCE          Mus musculus (house mouse)
ORGANISM        Mus musculus
REFERENCE       NIH-MGC http://mgc.nci.nih.gov/
AUTHORS        1 (bases 1 to 499)
JOURNAL        National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT        Unpublished (1999)
                Email: cgabbs-r@mail.nih.gov
                Tissue Procurement: Dr. James Lin, University of Iowa
                cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
                DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
                Clone Distribution: MGC clone distribution information can be
                found through the I.M.A.G.E. Consortium/LLNL at:
                http://image.llnl.gov
                This clone was contributed by the Brain Molecular Anatomy Project
                (BMAP)
                The following repetitive elements were found in this cDNA
                sequence: 3-108, >B1-FHSINE/Alu 122-195, >ID3#SINE/ID 333-363,
                >AT rich#Low_complexity 374-499, >B1_MM#SINE/Alu (matched
                complement)
                Seq primer: pyx-5.
                Location/Qualifiers
                1..499
                /organism="Mus musculus"
                /mol_type="mRNA"
                /strain="C57BL/6"
                /db_xref="taxon:10090"
                /clone="IMAGE:5719725"
                /tissue_type="whole brain"
                /dev_stage="embryo 15.5 dpc"
                /lab_host="DH10B (T1 phage resistant)"
                /clone_lib="NIH BMAP EX0"
                /note="Organ: brain; Vector: pyx-Asc; Site 1: EcoR I;
                Site 2: Not I; The library was constructed according to
                Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
                1996. Denatured mRNA was size fractionated on a 1% agarose
                gel. First strand cDNA synthesis was primed with an
                oligo-dT primer containing a Not I site. Double stranded
                cDNA was size selected according to mRNA size fraction.
                ligated with EcoR I adaptor, digested with Not I, and then
                cloned directionally into pyx-Asc vector. The library tag
                sequence located between the Not I site and the polyA
                tail, is GTCGCTGGAA. This library was created for the
                University of Iowa Mouse Brain Molecular Anatomy Project
                (BMAP). 'Gene Discovery in the Developing Mouse Nervous
                System', supported by National Institutes of Mental Health
                (NIMH), Hemin Chin, Ph.D., program coordinator."

FEATURES        source
                1..499
                /organism="Mus musculus"
                /mol_type="mRNA"
                /strain="C57BL/6J"
                /db_xref="taxon:10090"
                /clone="UI-M-CG0p-bnw-a-11-0-UI"
                /lab_host="DH10B (Life Technologies)"
                /clone_lib="NIH BMAP Ret4 S2"
                /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                polylinker; Site 1: Not I; Site 2: Eco RI; The
                NIH_BMAP Ret4_S2 library is a subtracted library,
                ultimately derived from mouse retina tissue libraries at
                various stages of development. For a detailed description
                of the library from which this clone was derived, please
                visit our web site at brainest.eng.uiowa.edu. The tissue
                for this library was contributed by Dr. Xin-Yuan Fu, Yale
                University School of Medicine"

ORIGIN
Query Match      100.0%; Score 25; DB 6; Length 499;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAAACACCACTGAGCTGGG 25
    |||||
Db 105 AAAAAACACCACTGAGCTGGG 129
    |||||

RESULT 5
BM935380/c
LOCUS          517 bp mRNA linear EST 29-APR-2002
DEFINITION    UI-M-CG0p-bnw-a-11-0-UI.r1 NIH_BMAP Ret4_S2 Mus musculus cDNA clone
                UI-M-CG0p-bnw-a-11-0-UI 5', mRNA sequence.
ACCESSION     BM935380
VERSION       BM935380.1 GI:19394532
KEYWORDS      EST.
SOURCE        Mus musculus (house mouse)
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarctomoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.
1 (bases 1 to 517)
Bonaldo,M.F., Lennon,G. and Soares,M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
889548
Contact: Chin, H
National Institute of Mental Health
6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
20892-9643, USA
Tel: 301 443 1706
Fax: 301 443 9890
Email: mEST@mail.nih.gov
Tissue Procurement: Dr. Xin-Yuan Fu, Yale University School of
Medicine
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).
The following repetitive elements were found in this cDNA
sequence: 41-114, >ID3#SINE/ID (matched complement) 128-283,
>B1_MM#SINE/Alu (matched complement) 296-367, >PB1D7#SINE/Alu
(matched complement)
Seq primer: M13 REVERSE.
                Location/Qualifiers
                1..517
                /organism="Mus musculus"
                /mol_type="mRNA"
                /strain="C57BL/6J"
                /db_xref="taxon:10090"
                /clone="UI-M-CG0p-bnw-a-11-0-UI"
                /lab_host="DH10B (Life Technologies)"
                /clone_lib="NIH BMAP Ret4 S2"
                /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                polylinker; Site 1: Not I; Site 2: Eco RI; The
                NIH_BMAP Ret4_S2 library is a subtracted library,
                ultimately derived from mouse retina tissue libraries at
                various stages of development. For a detailed description
                of the library from which this clone was derived, please
                visit our web site at brainest.eng.uiowa.edu. The tissue
                for this library was contributed by Dr. Xin-Yuan Fu, Yale
                University School of Medicine"

ORIGIN
Query Match      100.0%; Score 25; DB 3; Length 517;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAAACACCACTGAGCTGGG 25
    |||||
Db 131 AAAAAACACCACTGAGCTGGG 107
    |||||

RESULT 6
BB035891
LOCUS          664 bp mRNA linear EST 18-OCT-2001
DEFINITION    BB035891 RIKEN full-length enriched, 13 days embryo forelimb Mus
                musculus cDNA clone 5930406J05 3', mRNA sequence.
ACCESSION     BB035891
VERSION       BB035891.2 GI:16258342
KEYWORDS      EST.
SOURCE        Mus musculus (house mouse)
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Euarctomoglires; Glires; Rodentia;
                Sciurognathi; Muroidea; Muridae; Murinae; Mus.
1 (bases 1 to 664)
AraKawa,T., Carninci,P., Fukuda,S., Furuno,M., Hangsaki,T.,
Hara,A., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J.,
Konno,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K.,

```

TITLE
JOURNAL
COMMENT

Ohno, M., Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasuki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Takeda, Y., Tanaka, T., Toya, T., Muramatsu, M. and Hayashizaki, Y. RIKEN Mouse ESTs (Arakawa, T., et al. 2001) Unpublished (2001)

On Jun 10, 2000 this sequence version replaced gi:8442277.
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsr.riken.jp, URL:<http://genome.gsc.riken.jp/>
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. *Genome Res.* 10 (10), 1617-1630 (2000)
wagi, K., Fujiwake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer. *Genome Res.* 10 (11), 1757-1771 (2000)
Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara, Y. and Hayashizaki, Y.
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a non-redundant cDNA library. *Genome Res.* 11 (2), 281-289 (2001)
Kondo, S., Shinagawa, A., Saito, T., Kiyosawa, H., Yamanaka, I., Aizawa, K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K. and Hayashizaki, Y.
Computational Analysis of Full-length Mouse cDNAs Compared with Human Genome Sequences *Mamm. Genome.* 12, 673-677 (2001)
Please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

FEATURES
source

1. Location/Qualifiers
 - /organism="Mus musculus"
 - /mol_type="mRNA"
 - /strain="CS7BL/6J"
 - /db_xref="taxon:10090"
 - /clone="5930406J05"
 - /sex="mixed"
 - /tissue_type="forelimb"
 - /dev_stage="13 days embryo"
 - /lab_host="DH10B"
 - /clone_lib="RIKEN full-length enriched, 13 days embryo forelimb"
 - /note="Site_1: SalI; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGAGAGAAGATCCGAAGCTCTTTTTCCTTTTTTTTNN 3'], cDNA was prepared by using trihalose thermo-activated reverse transcriptase and subsequently enriched for full-length b cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 100.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGATTCTCGATTAAATAATACCCCCCCCCCCC 3']. cDNA was cloned into the XhoI and BamHI sites. Vector: a modified pBluescript KS(+) after bulk excision


```
RESULT 9
BY635162
LOCUS
DEFINITION
  BY635162 RIKEN full-length enriched, visual cortex Mus musculus
  cDNA clone K430066P18 3', mRNA sequence.
ACCESSION
  BY635162
VERSION
  BY635162.1 GI:26970344
KEYWORDS
  EST.
SOURCE
  Mus musculus (house mouse)
ORGANISM
  Mus musculus
  Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
  Sciurognathi; Muridea; Muridae; Murinae; Mus.
REFERENCE
  Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
  Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I.,
  Kiyosawa H., Yagi K., Tomaru Y., Hasegawa Y., Nogami A.,
  Schonbach C., Gojobori T., Baldarelli R., Hill D.P., Bult C.,
  Hume D.A., Quackenbush J., Schriml L.M., Kanapin A., Matsuda H.,
  Batalov S., Beisel K.W., Blake J.A., Bradt D., Brusic V.,
  Chothia C., Corbani L.E., Cousins S., Dalla E., Dragani T.A.,
  Fletcher C.F., Forrest A., Frazer K.S., Gaasterland T.,
  Gariboldi M., Gissi C., Godzik A., Gough J., Grimmond S.,
  Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D., Kanai A.,
  Kawai H., Kawasawa Y., Kedzierski R.M., King B.L., Konagaya A.,
  Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A., Maglott D.R.,
  Maltais L., Marchionni L., McKenzie L., Miki H., Nagashima T.,
  Numata K., Okido T., Favan W.J., Pertea G., Pesole G.,
  Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
  Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
  Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
  Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
  Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y.,
  Wells C., Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I.,
  Yang L., Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P.,
  Hayatsu N., Hirozane-Kishikawa T., Konno H., Nakamura M.,
  Sakazume N., Sato K., Shiraki T., Waki K., Kawai J., Aizawa K.,
  Arakawa T., Fukuda S., Hara A., Haseizume W., Imotani K., Ishii Y.,
  Itoh M., Kagawa I., Miyazaki A., Sakai K., Sasaki D., Shibata K.,
  Shingawa A., Yasunishi A., Yoshino M., Waterston R., Lander E.S.,
  Rogers J., Birney B. and Hayashizaki Y.
  Analysis of the mouse transcriptome based on functional annotation
  of 60,770 full-length cDNAs
  Nature 420, 563-573 (2002)
JOURNAL
  1246851
PUBMED
  Contact: Yoshihide Hayashizaki
  Laboratory for Genome Exploration Research Group, RIKEN Genomic
  Sciences Center (GSC), Yokohama Institute
  The Institute of Physical and Chemical Research (RIKEN)
  1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
  Tel: 81-45-503-9222
  Fax: 81-45-503-9216
  Email: genome-res@gscc.riken.jp, URL: http://genome.gsc.riken.jp/
  Aizawa K., Akimura T., Arakawa T., Carninci P., Fukuda S.,
  Hirozane T., Imotani K., Ishii Y., Itoh M., Kawai J., Konno H.,
  Miyazaki A., Murata M., Nakamura M., Nomura K., Numazaki R.,
  Ono M., Sakai K., Sakazume N., Sasaki D., Sato K., Shibata K.,
  Shiraki T., Tagami M., Waki K., Watahiki A., Muramatsu M. and
  Hayashizaki Y. Direct Submission
  Computational Analysis of Full-Length Mouse cDNAs Compared with
  Human Genome Sequences Mamm. Genome 12, 673-677 (2001)
  Normalization and subtraction of cap-trapper-selected cDNAs to
  prepare full-length cDNA libraries for rapid discovery of new
  genes. Genome Res. 10 (10), 1617-1630 (2000)
  RIKEN integrated sequence analysis (RISA) system--384-format
  sequencing pipeline with 384 multicapillary sequencer. Genome Res.
  10 (11), 1757-1771 (2000)
  Computer-based methods for the mouse full-length cDNA
  encyclopedia: real-time sequence clustering for construction of a
  nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
  cDNA library was prepared and sequenced in Mouse Genome
  Encyclopedia Project of Genome Exploration Research Group in Riken
  Genomic Sciences Center and Genome Science Laboratory in RIKEN.
```

Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Tissues were provided by Michela Fagiolini and Takao K. Hensch (
Laboratory for Neuronal Circuit Development Brain Science Institute
RIKEN 2-1 Hirosawa, Wako-shi, Saitama 351-0198 Japan) whose
assistance we gratefully acknowledge.
Please visit our web site (<http://genome.gsc.riken.go.jp>) for
further details.

FEATURES

Location/Qualifiers

1. .430

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="K430066P18"

/tissue_type="visual cortex"

/clone_lib="RIKEN full-length enriched, visual cortex"

ORIGIN

Query Match 93.6%; Score 23.4; DB 5; Length 430;

Best Local Similarity 96.0%; Pred. No. 15;

Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAAAACACCCACTGAGCTGGG 25

31 AAGAAACACCCACTGAGCTGGG 55

BF464082 450 bp mRNA linear EST 04-DEC-2000

UI-M-CGOp-bnw-a-11-0-UI.s1 NIH BMAP Ret4 S2 Mus musculus cDNA clone

UI-M-CGOp-bnw-a-11-0-UI 3', mRNA sequence.

BF464082

EST.

BF464082.1 GI:11533265

Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

Sciurognathi; Muridea; Muridae; Murinae; Mus.

1 (bases 1 to 450)

Bonaldo M.F., Lennon G. and Soares M.B.

Normalization and subtraction: two approaches to facilitate gene

discovery

Genome Res. 6 (9), 791-806 (1996)

889548

Contact: Chin, H

National Institute of Mental Health

6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD

20892-9643, USA

Tel: 301 443 1706

Fax: 301 443 9890

Email: mestr@mail.nih.gov

Oligo-dt track not found, Not I site shown in beginning of sequence

is likely internal to the message. cDNA library preparation: M.B.

Soares Lab Clone distribution: Researchers may obtain BMAP cDNA

clones from RESEARCH GENETICS. It should be noted that Bento Soares

is generating a small number of additional specialized

non-redundant arrays of BMAP cDNAs whose availability will be

considered under appropriate and limited collaborative arrangements

The following repetitive elements were found in this cDNA sequence:

159-230, >PBLD7#SINE/Alu 243-398, >B1_MM#SINE/Alu

Seq primer: M13 Forward

POLYA=No.

Location/Qualifiers

1. .450

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="UI-M-CGOp-bnw-a-11-0-UI"

/lab_host="DH10B (Life Technologies)"

/clone lib="NIH BMAP Ret4 S2"
 /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
 polylinker; Site 1: Not I; Site 2: Eco RI; The
 NIH_BMAP_Ret4_S2 library is a subtracted library,
 ultimately derived from mouse retina tissue libraries at
 various stages of development. For a detailed description
 of the library from which this clone was derived, please
 visit our web site at brainest.eng.uiowa.edu. The tissue
 for this library was contributed by Dr. Xin-Yuan Fu, Yale
 University School of Medicine
 TAG_SEQ=None found"

ORIGIN

Query Match 93.6%; Score 23.4; DB 2; Length 450;
 Best Local Similarity 96.0%; Pred. No. 16;
 Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAAACAACCACTGAGCTGGG 25
 |||||
 DB 395 ACAAAACAACCACTGAGCTGGG 419
 |||||

RESULT 11

BB760039 478 bp mRNA linear EST 17-OCT-2001
 BB760039 RIKEN full-length enriched, melanocyte Mus musculus cDNA
 clone G270124E08 3', mRNA sequence.

ACCESSION BB760039

VERSION BB760039.1 GI:16206348

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muroidea; Muridae; Murinae; Mus.

1 (bases 1 to 478)

REFERENCE

Akimura, T., Arakawa, T., Carninci, P., Furuno, M., Hanagaki, T.,
 Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Imotani, K.,
 Ishii, Y., Ito, M., Kawai, J., Kojima, Y., Konno, H., Kouda, M.,
 Matsuyama, T., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R.,
 Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sakazume, N.,
 Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,
 Sogabe, Y., Suzuki, H., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
 Tanaka, T., Tomaru, A., Toyota, T., Watahiki, A., Yasunishi, A.,
 Muramatsu, M., and Hayashizaki, Y.
 RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura, T., et al.
 2001)

TITLE

Unpublished (2001)
 Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216

Email: genome-res@gsc.riken.jp, URL: <http://genome.gsc.riken.jp/>
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
 Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new
 genes. Genome Res. 10 (10), 1617-1630 (2000)
 wagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,
 Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T.,
 Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A.,
 and Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10
 (11), 1757-1771 (2000)

Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P.,
 Sugahara, Y., and Hayashizaki, Y.
 Computer-based methods for the mouse full-length cDNA
 encyclopedia: real-time sequence clustering for construction of a
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

Please visit our web site (<http://genome.gsc.riken.go.jp>) for
 further details.

FEATURES

source

Location/Qualifiers
 1..478
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="G270124E08"
 /cell_type="melanocyte"
 /clone_lib="RIKEN full-length enriched, melanocyte"

ORIGIN

Query Match 93.6%; Score 23.4; DB 2; Length 478;
 Best Local Similarity 96.0%; Pred. No. 16;
 Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAAACAACCACTGAGCTGGG 25
 |||||
 DB 79 AAGAAACAACCACTGAGCTGGG 103
 |||||

RESULT 12

BY650318 395 bp mRNA linear EST 16-DEC-2002
 BY650318 RIKEN full-length enriched, visual cortex Mus musculus
 cDNA clone K530047N21 3', mRNA sequence.

ACCESSION BY650318

VERSION BY650318.1 GI:27007288

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muroidea; Muridae; Murinae; Mus.

1 (bases 1 to 395)

REFERENCE

Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
 Nikaido, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I.,
 Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,
 Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D. P., Bult, C.,
 Hume, D. A., Quackenbush, J., Schriml, L. M., Kanapin, A., Matsuda, H.,
 Batalov, S., Beisel, K. W., Blake, J. A., Bradt, D., Brusic, V.,
 Chothia, C., Corbani, L. E., Cousins, S., Dalla, E., Dragani, T. A.,
 Fletcher, C. F., Forrest, A., Frazer, K. S., Gaasterland, T.,
 Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S.,
 Gustincich, S., Hirokawa, N., Jackson, I. J., Jarvis, E. D., Kanai, A.,
 Kawaji, H., Kawasawa, Y., Kedzierski, R. M., King, B. L., Konagaya, A.,
 Kurochkin, I. V., Lee, Y., Lenhard, B., Lyons, P. A., Maglott, D. R.,
 Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T.,
 Numata, K., Okido, T., Pavan, W. J., Pertea, G., Pesole, G.,
 Petrovsky, N., Pillai, R., Pontius, J. U., Qi, D., Ramachandran, S.,
 Ravasi, T., Reed, J. C., Reed, D. J., Reid, J., Ring, B. Z., Ringwald, M.,
 Sandelin, A., Schneider, C., Semple, C. A., Setou, M., Shimada, K.,
 Sultana, R., Takenaka, Y., Taylor, M. S., Teasdale, R. D., Tomita, M.,
 Vekardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y.,
 Wells, C., Wilming, L. G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I.,
 Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P.,
 Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M.,
 Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,
 Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y.,
 Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K.,
 Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E. S.,
 Rogers, J., Birney, E., and Hayashizaki, Y.

Analysis of the mouse transcriptome based on functional annotation
 of 60,770 full-length cDNAs

JOURNAL

Nature 420, 563-573 (2002)

PUBMED

12466851

COMMENT

Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-res@genome.riken.jp, URL: <http://genome.gsc.riken.jp/>
 Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission
 Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
 Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
 RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)
 Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
 Tissues were provided by Michela Fagiolini and Takao K. Hensch (Laboratory for Neuronal Circuit Development Brain Science Institute RIKEN 2-1 Hirotsawa, Wako-shi, Saitama 351-0198 Japan) whose assistance we gratefully acknowledge.
 Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES

source

Location/Qualifiers
 1. 395
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="K530047N21"
 /tissue_type="visual cortex"
 /clone_lib="RIKEN full-length enriched, visual cortex"

ORIGIN

Query Match 88.0%; Score 22; DB 5; Length 395;
 Best Local Similarity 100.0%; Pred. No. 65;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 AACACACCACCTGAGCTGGG 25

|||||
 1 AACACACCACCTGAGCTGGG 22

RESULT 13

BY674654

LOCUS

DEFINITION
 BY674654 RIKEN full-length enriched, 14.5 days embryo df/df
 Rathke's pouches Mus musculus cDNA clone K820041010 3', mRNA
 sequence.

ACCESSION

VERSION

BY674654.1

KEYWORDS

SOURCE

EST.

Mus musculus (house mouse)

ORGANISM

Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

AUTHORS

Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
 Kikado, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I.,
 Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,
 Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C.,
 Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H.,
 Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusis, V.,
 Chothia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A.,

Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S., Gunatich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R., Maltais, L., Marchionni, L., Mckenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W.J., Pertea, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S., Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M., Sandelin, A., Schneider, C., Semple, C.A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomika, M., Verardo, C., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, R.C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S., Rogers, J., Birney, E. and Hayashizaki, Y.

TITLE

JOURNAL

PUBMED

COMMENT

12456851

Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic

Sciences Center (GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-res@genome.riken.jp, URL: <http://genome.gsc.riken.jp/>

Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S.,

Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H.,

Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R.,

Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K.,

Shiraki, T., Tagami, M., Waki, K., Watahiki, A., Muramatsu, M. and

Hayashizaki, Y. Direct Submission

Computational Analysis of Full-Length Mouse cDNAs Compared with

Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new

genes. Genome Res. 10 (10), 1617-1630 (2000)

RIKEN integrated sequence analysis (RISA) system--384-format

sequencing pipeline with 384 multicapillary sequencer. Genome Res.

10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA

encyclopedia: real-time sequence clustering for construction of a

nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

cDNA library was prepared and sequenced in Mouse Genome

Encyclopedia Project of Genome Exploration Research Group in Riken

Genomic Sciences Center and Genome Science Laboratory in Riken.

Division of Experimental Animal Research in Riken contributed to

prepare mouse tissues.

Tissues were provided by Michelle Brinkmeier and Sally Camper (

Dept. Human Genetics University of Michigan Medical School 4301

MSRB 3 1500 W. Medical Center Dr. Ann Arbor, MI 48109-0638 USA)

whose assistance we gratefully acknowledge.

Please visit our web site (<http://genome.gsc.riken.go.jp>) for

further details.

Location/Qualifiers

1. 394

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

/clone="K820041010"

/tissue_type="Rathke's pouches"

/dev_stage="14.5 days embryo df/df"

/clone_lib="RIKEN full-length enriched, 14.5 days embryo

df/df Rathke's pouches"

ORIGIN

Query Match

84.0%; Score 21; DB 5; Length 394;

Best Local Similarity 100.0%; Pred. No. 1.8e+02; Mismatches 0; Indels 0; Gaps 0;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 5 AACAAACCACTGAGCTGGG 25
 |||||
 Db 1 AACAAACCACTGAGCTGGG 21
 |||||

RESULT 14
 CV525379
 LOCUS
 DEFINITION
 Mdlv4011c18.y1 Mdlv Malus x domestica cDNA clone Mdlv4011c18 5',
 similar to TR:Q9ZNS3 Q9ZNS3 RIBOSOMAL PROTEIN S27. ;, mRNA
 sequence.

ACCESSION
 VERSION
 KEYWORDS
 SOURCE

ORGANISM

Malus x domestica
 Malus x domestica
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 rosids; eurosids 1; Rosales; Rosaceae; Maloideae; Malus.

REFERENCE
 AUTHORS
 1 (bases 1 to 172)
 Korbán,S., Vodkin,L., Liu,L., Gasic,K., Gonzales,O., Hernandez,A.,
 Aldwinckle,H., Malnoy,M., Carroll,N., Goldsbrough,P., Orvis,K.,
 Clifton,S., Pape,D., Marra,M., Hillier,L., Martin,J., Wylie,T.,
 Dante,M., Theising,B., Bowers,Y., Gibbons,M., Ritter,E., Ronko,I.,
 Tsagarisvili,R., Kennedy,S., Waterston,R. and Wilson,R.,
 Apple Functional Genomics grant - NSF 0321702
 Unpublished (2004)

TITLE
 JOURNAL

COMMENT

Contact: Schuyler S. Korban
 Apple Functional Genomics grant - NSF 0321702
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu

Library materials provided by: Schuyler S. Korban Library
 constructed by: D.O.Gonzalez/L. Vodkin Library sequenced by:
 Washington University Genome Sequencing Center
 WashU EST name: aan94b09.y1

Trace considered overall poor quality
 Possible reversed clone: similarity on wrong strand
 Seq primer: -40RP from Gibco
 High quality sequence stop: 1.

FEATURES
 source

Location/Qualifiers
 1..172
 /organism="Malus x domestica"
 /mol_type="mRNA"
 /cultivar="GoldRush"
 /db_xref="taxon:3750"
 /clone="Mdlv4011c18"
 /tissue_type="Leaf"
 /lab_host="DH10B ampicillin resistant"
 /clone_lib="Mdlv"
 /note="Vector: pSPORT 1; Site 1: NotI; Site 2: Sal I;
 Total RNA was extracted from freeze dried leaf tissue
 Stage I, using a standard Phenol chloroform extraction
 method. Poly(A)+mRNA was isolated from total RNA using the
 PolyAtract mRNA Isolation system III (Promega). The
 library was prepared using the Invitrogen Life
 Technologies, Superscript Plasmid System with gateway
 technology for cDNA Synthesis and cloning. Complementary
 DNA was synthesized from mRNA using an anchored Poly (dT)
 sequence with a NotI restriction site. Sal I linker
 adapters were ligated to the blunt ended cDNA fragments
 followed by restriction with NotI. The cDNA fragments were
 directionally cloned into the NotI-SalI restriction site
 of the pSPORT 1 vector. The ligated cDNA fragments were
 transformed into E. coli ElectroMax DH10B host cells.
 Transformation efficiency: 8.5E+07 Average insert Size by
 PCR: 1100 bp"

Query Match 80.8%; Score 20.2; DB 7; Length 172;
 Best Local Similarity 88.0%; Pred. No. 3.9e+02;
 Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 Qy 1 AAAAAACAACCACTGAGCTGGG 25
 |||||
 Db 12 AAAAAACAACCACTGAGCTGGG 36
 |||||

RESULT 15
 CV525125
 LOCUS

DEFINITION

CV525125
 Mdlv4010g21.y1 Mdlv Malus x domestica cDNA clone Mdlv4010g21 5',
 similar to TR:Q9ZNS3 Q9ZNS3 RIBOSOMAL PROTEIN S27. ;, mRNA
 sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Malus x domestica
 Malus x domestica
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 rosids; eurosids 1; Rosales; Rosaceae; Maloideae; Malus.

REFERENCE

AUTHORS

1 (bases 1 to 256)
 Korbán,S., Vodkin,L., Liu,L., Gasic,K., Gonzales,O., Hernandez,A.,
 Aldwinckle,H., Malnoy,M., Carroll,N., Goldsbrough,P., Orvis,K.,
 Clifton,S., Pape,D., Marra,M., Hillier,L., Martin,J., Wylie,T.,
 Dante,M., Theising,B., Bowers,Y., Gibbons,M., Ritter,E., Ronko,I.,
 Tsagarisvili,R., Kennedy,S., Waterston,R. and Wilson,R.,
 Apple Functional Genomics grant - NSF 0321702
 Unpublished (2004)

TITLE
 JOURNAL

COMMENT

Contact: Schuyler S. Korban
 Apple Functional Genomics grant - NSF 0321702
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu

Library materials provided by: Schuyler S. Korban Library
 constructed by: D.O.Gonzalez/L. Vodkin Library sequenced by:
 Washington University Genome Sequencing Center
 WashU EST name: aan89d11.y1

Trace considered overall poor quality
 Possible reversed clone: similarity on wrong strand
 Seq primer: -40RP from Gibco
 High quality sequence stop: 1.
 Location/Qualifiers

FEATURES
 source

1..256
 /organism="Malus x domestica"
 /mol_type="mRNA"
 /cultivar="GoldRush"
 /db_xref="taxon:3750"
 /clone="Mdlv4010g21"
 /tissue_type="Leaf"
 /lab_host="DH10B ampicillin resistant"
 /clone_lib="Mdlv"
 /note="Vector: pSPORT 1; Site 1: NotI; Site 2: Sal I;
 Total RNA was extracted from freeze dried leaf tissue
 Stage I, using a standard Phenol chloroform extraction
 method. Poly(A)+mRNA was isolated from total RNA using the
 PolyAtract mRNA Isolation system III (Promega). The
 library was prepared using the Invitrogen Life
 Technologies, Superscript Plasmid System with gateway
 technology for cDNA Synthesis and cloning. Complementary
 DNA was synthesized from mRNA using an anchored Poly (dT)
 sequence with a NotI restriction site. Sal I linker
 adapters were ligated to the blunt ended cDNA fragments
 followed by restriction with NotI. The cDNA fragments were
 directionally cloned into the NotI-SalI restriction site
 of the pSPORT 1 vector. The ligated cDNA fragments were
 transformed into E. coli ElectroMax DH10B host cells.
 Transformation efficiency: 8.5E+07 Average insert Size by

```

ORIGIN
Query Match      80.8%; Score 20.2; DB 7; Length 256;
Best Local Similarity 88.0%; Pred. No. 4e+02; 3; Indels 0; Gaps 0;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAACACACCACCTGAGCTGGG 25
    ||||| ||||| ||||| ||||| |||||
Db 15 AAAAAACACACCACCTGAGCTGGG 39
    ||||| ||||| ||||| ||||| |||||

RESULT 16
BX764584
LOCUS
DEFINITION
Xenopus tropicalis (western clawed frog)
mRNA sequence.
ACCESSION
BX764584
VERSION
BX764584.1 GI:39671793
KEYWORDS
Xenopus tropicalis (western clawed frog)
SOURCE
Xenopus tropicalis
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
Xenopodinae; Xenopus; Silurana.
REFERENCE
1 (bases 1 to 375)
Croning M.D.R., Ashurst J.L., Taylor R., Zorn A.M. and Rogers J.
Sanger Xenopus tropicalis EST project 2001 (11_2003)
TITLE
Xenopus tropicalis EST project 2001 (11_2003)
JOURNAL
Unpublished (2003)
COMMENT
Contact: Croning MDR
Hinxton, Cambridgeshire, CB10 1SA, UK
Email: trop@sanger.ac.uk
This sequence is from a Xenopus Gene Collection (XGC) library
constructed by Aaron M. Zorn.
cdna was oligo dt primed from sug of poly A+ RNA from stages 10-13
gastrulae. EcoRI-NotI cut cDNA was then ligated into pCS107 with
EcoRI at the 5' end and NotI at the 3' end.
Vector: pCS107; Site 1: EcoRI; Site 2: NotI
Host: Escherichia coli XL1-blue
Sanger Xenopus tropicalis EST project 2001
TROPICALIS_SEQUENCE_ID: Tgas140j16.q1kT7
Sequencing primer: T7
This sequence is from a Xenopus Gene Collection (XGC) library
constructed by Aaron M. Zorn.
cdna was oligo dt primed from sug of poly A+ RNA from stages 10-13
gastrulae. EcoRI-NotI cut cDNA was then ligated into pCS107 with
EcoRI at the 5' end and NotI at the 3' end.
Vector: pCS107; Site 1: EcoRI; Site 2: NotI
Host: Escherichia coli XL1-blue.
FEATURES
source
Location/Qualifiers
1..375
/organism="Xenopus tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="Tgas140j16"
/dev_stage="gastrula (stages 10.5-12 mixed)"
/lab_host="Escherichia coli XL1-blue"
/clone_lib="XGC-gastrula"
/notes="Vector: pCS107; Site 1: EcoRI; Site 2: NotI; cdna
was oligo dt primed from sug of poly A+ RNA from stages
10-13 gastrulae. EcoRI-NotI cut cDNA was then ligated
into pCS107 with EcoRI at the 5' end and NotI at the 3'
end."

ORIGIN
Query Match      80.8%; Score 20.2; DB 1; Length 577;
Best Local Similarity 88.0%; Pred. No. 4.2e+02;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAACACACCACCTGAGCTGGG 25
    ||||| ||||| ||||| ||||| |||||
Db 138 AAAAAACACACCACCTGAGCTGGG 114
    ||||| ||||| ||||| ||||| |||||

RESULT 18
BX746867
LOCUS
DEFINITION
mRNA sequence.
ACCESSION
BX746867
VERSION
BX746867.1 GI:38419607
KEYWORDS
Xenopus tropicalis (western clawed frog)
SOURCE
Xenopus tropicalis
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
Xenopodinae; Xenopus; Silurana.
REFERENCE
1 (bases 1 to 596)
Croning M.D.R., Ashurst J.L., Taylor R., Zorn A.M. and Rogers J.
Sanger Xenopus tropicalis EST project 2001 (11_2003)
TITLE
Xenopus tropicalis EST project 2001 (11_2003)
JOURNAL
Unpublished (2003)
COMMENT
Contact: Croning MDR
Sanger Institute
Hinxton, Cambridgeshire, CB10 1SA, UK
Email: trop@sanger.ac.uk

```

Sanger Xenopus tropicalis EST project 2001
TROPICALIS SEQUENCE ID: TGas066n23.q1kt7
Sequencing primer: T7
This sequence is from a Xenopus Gene Collection (XGC) library
constructed by Aaron M. Zorn.
cDNA was oligo dt primed from sug of poly A+ RNA from stages 10-13
gastrulae. EcoRI-NotI cut cDNA was then ligated into pCS107 with
EcoRI at the 5' end and NotI at the 3' end.
Vector: pCS107; Site 1: EcoRI; Site 2: NotI
Host: Escherichia coli XL1-blue.

FEATURES

source

```

1..596
    /organism="Xenopus tropicalis"
    /mol_type="mRNA"
    /db_xref="taxon:8364"
    /clone="TGas066n23"
    /dev_stage="gastrula (stages 10.5-12 mixed)"
    /lab_host="Escherichia coli XL1-blue"
    /clone_lib="XGC-gastrula"
    /note="Vector: pCS107; Site 1: EcoRI; Site 2: NotI; cDNA
was oligo dt primed from sug of poly A+ RNA from stages
10-13 gastrulae. EcoRI-NotI cut cDNA was then ligated
into pCS107 with EcoRI at the 5' end and NotI at the 3'
end."

```

ORIGIN

```

Query Match      80.8%; Score 20.2; DB 5; Length 596;
Best Local Similarity 88.0%; Pred. No. 4.2e+02;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

QY 1 AAAAAACACCACTGAGCTGGG 25

```
|||||
|||||
|||||
|||||
|||||

```

Db 260 AAAAAACCACTGACTGGG 284

RESULT 19

BQ526123/c

LOCUS

```

DEFINITION      BQ526123 600 bp mRNA linear EST 10-JUN-2002
IMAGE:5381083 5', mRNA sequence.

```

ACCESSION

BQ526123

VERSION

BQ526123.1

GI:21384992

KEYWORDS

EST.

SOURCE

Xenopus tropicalis (western clawed frog)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;

Xenopodinae; Xenopus; Silurana.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

NIH-XGC <http://image.llnl.gov/image/html/xenopuslib.info.shtml>.
National Institute of Child Health and Human Development, National
Cancer Institute, Xenopus Gene Collection
Unpublished (2002)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
CDNA Library Preparation:
CDNA Library Arrayed by: The I.M.A.G.E. Consortium/LLNL
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC)
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
info.image.llnl.gov

Plate: LLAM11970 row: I column: 20

Seq primer: M13RP1 reverse primer (ABI).

FEATURES

source

```

1..600
    /organism="Xenopus tropicalis"
    /mol_type="mRNA"
    /db_xref="taxon:8364"
    /clone="IMAGE:5381083"
    /tissue_type="cadpole"
    /dev_stage="embryo, stages 40-45"
    /lab_host="DH10B (phage-resistant)"

```

/clone lib="NICHD XGC Emb8"
/note="Vector: pCMV-SPORT6.1; Site 1: NotI; Site 2: EcoRV;
Cloned unidirectionally. Primer: Oligo dt. Average insert
size 2.1 kb. Constructed by Invitrogen. Note: This is a
Xenopus Gene Collection (XGC) library."

ORIGIN

```

Query Match      80.8%; Score 20.2; DB 5; Length 600;
Best Local Similarity 88.0%; Pred. No. 4.2e+02;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

QY 1 AAAAAACACCACTGAGCTGGG 25

```
|||||
|||||
|||||
|||||
|||||

```

Db 258 AAAAAACCACTGTACTGGG 234

RESULT 20

AL681607/c

LOCUS

```

DEFINITION      AL681607 606 bp mRNA linear EST 10-NOV-2003
mRNA sequence.

```

ACCESSION

AL681607

VERSION

AL681607.2

GI:38253507

KEYWORDS

EST.

SOURCE

Xenopus tropicalis (western clawed frog)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;

Xenopodinae; Xenopus; Silurana.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

On Mar 18, 2002 this sequence version replaced gi:19537981.
Contact: Taylor R
Sanger Institute
Hinxton, Cambridgeshire, CB10 1SA, UK
Email: trop@sanger.ac.uk
This sequence is from a Xenopus Gene Collection (XGC) library
constructed by Aaron M. Zorn.
cDNA was oligo dt primed from sug of poly A+ RNA from stages 10-13
gastrulae. EcoRI-NotI cut cDNA was then ligated into pCS107 with
EcoRI at the 5' end and NotI at the 3' end.
Vector: pCS107; Site 1: EcoRI; Site 2: NotI
Host: Escherichia coli XL1-blue

Sanger Xenopus tropicalis EST project 2001

TROPICALIS SEQUENCE ID: TGas066n23.plcSP6

Sequencing primer: SP6.

Location/Qualifiers

1..606

/organism="Xenopus tropicalis"

/mol_type="mRNA"

/db_xref="taxon:8364"

/clone="TGas066n23"

/dev_stage="gastrula (stages 10.5-12 mixed)"

/lab_host="Escherichia coli XL1-blue"

/clone_lib="XGC-gastrula"

/note="Vector: pCS107; Site 1: EcoRI; Site 2: NotI; cDNA

was oligo dt primed from sug of poly A+ RNA from stages

10-13 gastrulae. EcoRI-NotI cut cDNA was then ligated

into pCS107 with EcoRI at the 5' end and NotI at the 3'

end."

ORIGIN

```

Query Match      80.8%; Score 20.2; DB 1; Length 606;
Best Local Similarity 88.0%; Pred. No. 4.2e+02;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

QY 1 AAAAAACACCACTGAGCTGGG 25

```
|||||
|||||
|||||
|||||
|||||

```

Db 356 AAAAAACCACTGTACTGGG 332

```

RESULT 21
AL860221/c
LOCUS
DEFINITION
AL860221 XGC-egg Xenopus tropicalis cDNA clone TEGg071011 5', mRNA
sequence.
ACCESSION
AL860221
VERSION
AL860221.2 GI:38655576
KEYWORDS
SOURCE
Xenopus tropicalis (western clawed frog)
ORGANISM
Xenopus tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
Xenopodinae; Xenopus; Silurana.
1 (bases 1 to 653)
Croning,M.D.R., Ashurst,J.L., Taylor,R., Zorn,A.M. and Rogers,J.
Sanger Xenopus tropicalis EST project 2001 (11_2003)
UNPUBLISHED (2003)
COMMENT
On Sep 15, 2002 this sequence version replaced gi:22880402.
Contact: Taylor R
Sanger Institute
Hinxton, Cambridgeshire, CB10 1SA, UK
Email: trop@sanger.ac.uk
TROPICALIS_SEQUENCE_ID: TEGg071011.plkSP6
Sequencing primer: SP6
This sequence is from a Xenopus Gene Collection (XGC) library
constructed by Aaron M. Zorn.
cDNA was oligo dt primed from Sug of poly A+ RNA from egg.
EcoRI-NotI cut cDNA was then ligated into pCS107 with EcoRI at the
5' end and NotI at the 3' end.
Vector: pCS107; Site_1: EcoRI; Site_2: NotI
Host: Escherichia coli XL1-blue.
Location/Qualifiers
1..653
/organism="Xenopus tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="TEGg071011"
/dev_stage="egg"
/lab_host="Escherichia coli XL1-blue"
/clone_lib="XGC-egg"
/note="Vector: pCS107; Site_1: EcoRI; Site_2: NotI; cDNA
was oligo dt primed from Sug of poly A+ RNA from egg.
EcoRI-NotI cut cDNA was then ligated into pCS107 with
EcoRI at the 5' end and NotI at the 3' end"
ORIGIN
Query Match 80.8%; Score 20.2; DB 1; Length 653;
Best Local Similarity 88.0%; Pred. No. 4.3e+02;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAACACCACCTGAGCTGGG 25
||||| ||||| ||||| ||||| |||||
DB 169 AAAAAACACCACCTGAGCTGGG 145
||||| ||||| ||||| ||||| |||||

RESULT 22
AL854589/c
LOCUS
DEFINITION
AL854589 XGC-egg Xenopus tropicalis cDNA clone TEGg020p23 5', mRNA
sequence.
ACCESSION
AL854589
VERSION
AL854589.2 GI:38630096
KEYWORDS
SOURCE
Xenopus tropicalis (western clawed frog)
ORGANISM
Xenopus tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
Xenopodinae; Xenopus; Silurana.
1 (bases 1 to 654)
Croning,M.D.R., Ashurst,J.L., Taylor,R., Zorn,A.M. and Rogers,J.
Sanger Xenopus tropicalis EST project 2001 (11_2003)
UNPUBLISHED (2003)
COMMENT
On Sep 15, 2002 this sequence version replaced gi:22874809.
Contact: Taylor R
Sanger Institute
Hinxton, Cambridgeshire, CB10 1SA, UK
Email: trop@sanger.ac.uk
TROPICALIS_SEQUENCE_ID: TEGg020p23.plkSP6
Sequencing primer: SP6
This sequence is from a Xenopus Gene Collection (XGC) library
constructed by Aaron M. Zorn.
cDNA was oligo dt primed from Sug of poly A+ RNA from egg.
EcoRI-NotI cut cDNA was then ligated into pCS107 with EcoRI at the
5' end and NotI at the 3' end.
Vector: pCS107; Site_1: EcoRI; Site_2: NotI
Host: Escherichia coli XL1-blue.
Location/Qualifiers
1..654
/organism="Xenopus tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="TEGg020p23"
/dev_stage="egg"
/lab_host="Escherichia coli XL1-blue"
/clone_lib="XGC-egg"
/note="Vector: pCS107; Site_1: EcoRI; Site_2: NotI; cDNA
was oligo dt primed from Sug of poly A+ RNA from egg.
EcoRI-NotI cut cDNA was then ligated into pCS107 with
EcoRI at the 5' end and NotI at the 3' end"
ORIGIN
Query Match 80.8%; Score 20.2; DB 1; Length 654;
Best Local Similarity 88.0%; Pred. No. 4.3e+02;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAACACCACCTGAGCTGGG 25
||||| ||||| ||||| ||||| |||||
DB 359 AAAAAACACCACCTGAGCTGGG 335
||||| ||||| ||||| ||||| |||||

RESULT 23
CR584874/c
LOCUS
DEFINITION
CR584874 XGC-tailbud-head Xenopus tropicalis cDNA clone THdA045d02
5', mRNA sequence.
ACCESSION
CR584874
VERSION
CR584874.1 GI:50584874
KEYWORDS
SOURCE
Xenopus tropicalis (western clawed frog)
ORGANISM
Xenopus tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
Xenopodinae; Xenopus; Silurana.
1 (bases 1 to 654)
Croning,M.D.R., Ashurst,J.L., Taylor,R., Garrett,N. and Rogers,J.
Sanger Xenopus tropicalis EST project 2001 (2004)
UNPUBLISHED (2004)
COMMENT
Contact: Croning MDR
Sanger Institute
Hinxton, Cambridgeshire, CB10 1SA, UK
Email: trop@sanger.ac.uk
TROPICALIS_SEQUENCE_ID: THdA045d02.plkSP6
This sequence is from a Xenopus Gene Collection (XGC) library
constructed by Nigel Garrett.
Seq primer: SP6.
Location/Qualifiers
1..654
/organism="Xenopus tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="THdA045d02"
/dev_stage="tailbud head (stage 28-30)"
/lab_host="Escherichia coli DH10B."
FEATURES
source

```

/clone.lib="XGC-tailbud-head"
 /note="Vector: pCS107; Site 1: EcoRI; Site 2: NotI; cDNA
 was oligo dt primed from Sug of poly A+ RNA from tailbud
 head. EcoRI-NotI cut cDNA was then ligated into pCS107
 with EcoRI at the 5' end and NotI at the 3' end."

ORIGIN

Query Match 80.8%; Score 20.2; DB 7; Length 654;
 Best Local Similarity 88.0%; Pred. No. 4.3e+02;
 Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAACACCACTGAGCTGGG 25
 |||||

Db 97 AAAAAACCACTGACTGGG 73
 |||||

RESULT 24

AL885471/c
 LOCUS AL885471 XGC-egg Xenopus tropicalis cDNA clone TEGG044013 5', mRNA
 DEFINITION
 AL885471
 VERSION
 AL885471.2 GI:38676588
 KEYWORDS
 EST
 ORGANISM
 Xenopus tropicalis (western clawed frog)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Xenopodinae; Xenopus; Silurana.

REFERENCE

1 (bases 1 to 656)
 Croning, M.D.R., Ashurst, J.L., Taylor, R., Zorn, A.M. and Rogers, J.
 Sanger Xenopus tropicalis EST project 2001 (11_2003)
 Unpublished (2003)
 On Sep 16, 2002 this sequence version replaced gi:22926745.

CONTACT: Taylor R

Sanger Institute

Hinxton, Cambridgeshire, CB10 1SA, UK

Email: trop@sanger.ac.uk

Sanger Xenopus tropicalis EST project 2001

TROPICALIS SEQUENCE ID: TEGG044013.p1kSP6

Sequencing primer: SP6

This sequence is from a Xenopus Gene Collection (XGC) library

constructed by Aaron M. Zorn.

cDNA was oligo dt primed from Sug of poly A+ RNA from egg.

EcoRI-NotI cut cDNA was then ligated into pCS107 with EcoRI at the

5' end and NotI at the 3' end.

Vector: pCS107; Site 1: EcoRI; Site 2: NotI

Host: Escherichia coli XL1-blue.

FEATURES

source

1..656
 Location/Qualifiers
 /organism="Xenopus tropicalis"
 /mol_type="mRNA"
 /db_xref="taxon:8364"
 /clone="TEGG044013"
 /dev_stage="egg"
 /lab_host="Escherichia coli XL1-blue"
 /clone_lib="XGC-egg"
 /note="Vector: pCS107; Site 1: EcoRI; Site 2: NotI; cDNA
 was oligo dt primed from Sug of poly A+ RNA from egg.
 EcoRI-NotI cut cDNA was then ligated into pCS107 with
 EcoRI at the 5' end and NotI at the 3' end"

ORIGIN

Query Match 80.8%; Score 20.2; DB 1; Length 656;
 Best Local Similarity 88.0%; Pred. No. 4.3e+02;
 Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAACACCACTGAGCTGGG 25
 |||||

Db 518 AAAAAACCACTGACTGGG 494
 |||||

RESULT 25

CX359039/c

LOCUS CX359039 719 bp mRNA linear EST 05-JAN-2005
 DEFINITION JGI_XZT410.fwd NIH XGC troptad5 Xenopus tropicalis cDNA clone
 IMAGE:7580481 5', mRNA sequence.

ACCESSION

VERSION CX359039

KEYWORDS

SOURCE EST.

ORGANISM

Xenopus tropicalis (western clawed frog)
 Xenopus tropicalis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Xenopodinae; Xenopus; Silurana.

REFERENCE

AUTHORS 1 (bases 1 to 719)

Richardson, P., Lucas, S., Rokhe, D., Dettler, J.C., Ng, D.C.,

Brokstein, P. and Lindquist, E.A.

DOE Joint Genome Institute

Unpublished (2004)

Other ESTs: JGI_XZT410.rev

Contact: Lindquist, E.A., Richardson, P.

DOE Joint Genome Institute

2800 Mitchell Drive, Walnut Creek, CA 94598, USA

Tel: 925 296 5600

Fax: 925 296 5710

Email: cdna@jgi-psf.org

Tissue Procurement: Richard M. Harland Laboratory, University of

California, Berkeley; http://tropicalis.berkeley.edu/home

cDNA library Preparation: Richard M. Harland Laboratory, University

of California, Berkeley

DNA Sequencing: DOE Joint Genome Institute; http://www.jgi.doe.gov

Clone Distribution: I.M.A.G.E. Consortium/LINL;

http://image.llnl.gov

Naming Conventions: EST name is generated by the concatenation of

the JGI Clone ID and the direction of sequencing. The suffix '.fwd'

indicates a forward sequencing read of the insert. It does not

necessarily reflect the orientation of the insert.

Plate: XZT 0005 row: c column: 7

High quality sequence stop: 719.

FEATURES

source

1..719
 Location/Qualifiers
 /organism="Xenopus tropicalis"
 /mol_type="mRNA"
 /db_xref="taxon:8364"
 /clone="IMAGE:7580481"
 /tissue_type="whole embryo"
 /dev_stage="Tadpole (st. 36-41)"
 /lab_host="E. coli XL1-Blue derivative, Stratagene
 Electropen-Blue"
 /clone_lib="NIH XGC troptad5"
 /note="Vector: pCS108; Site 1: Sall; Site 2: NotI; Tadpole
 library constructed by Russell B. Fletcher in R. Harland's
 lab using poly A RNA and oligo dt primers (Invitrogen
 Superscript Plasmid System for cDNA Synthesis and
 Cloning). Sall (5' end) -NotI (3' end) cDNA was inserted
 into vector pCS108
 (http://mcb.berkeley.edu/labs/harland/pages/plasmids.html)"

ORIGIN

Query Match 80.8%; Score 20.2; DB 8; Length 719;
 Best Local Similarity 88.0%; Pred. No. 4.3e+02;
 Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAACACCACTGAGCTGGG 25
 |||||

Db 558 AAAAAACCACTGTACTGGG 534
 |||||

RESULT 26

CX408645/c

LOCUS CX408645

DEFINITION JGI_XZT31734.fwd NIH XGC troptad5 Xenopus tropicalis cDNA clone

IMAGE:7607944 5', mRNA sequence.

ACCESSION CX408645

744 bp mRNA linear EST 06-JAN-2005
 JGI_XZT31734.fwd NIH XGC troptad5 Xenopus tropicalis cDNA clone
 IMAGE:7607944 5', mRNA sequence.

of California, Berkeley
 DNA Sequencing: DOE Joint Genome Institute: <http://www.jgi.doe.gov>
 Clone Distribution: I.M.A.G.E. Consortium/LLNL:
<http://image.llnl.gov>
 Naming Conventions: EST name is generated by the concatenation of
 the JGI Clone Id and the direction of sequencing. The suffix '.fwd'
 indicates a forward sequencing read of the insert. It does not
 necessarily reflect the orientation of the insert.
 Plate: XZG 0429 row: j column: 18
 High quality sequence stop: 755.

FEATURES

source

```

1..761
Location/Qualifiers
  1..761
  /organism="Xenopus tropicalis"
  /mol_type="mRNA"
  /db_xref="taxon:8364"
  /clone="IMAGE:7556468"
  /tissue_type="whole embryo"
  /dev_stage="Gastrula (st. 10.5-12.5)"
  /lab_host="E. coli XL1-Blue derivative, Stratagene
Electrofen-Blue"
  /clone_lib="NIH_XGC_tropGas7"
  /note="Vector: PCS108; Site 1: SalI; Site 2: NotI;
Gastrula library constructed by Russell B. Fletcher in R.
Harland's lab using poly A RNA and oligo dt primers
(Invitrogen SuperScript Plasmid System for cDNA Synthesis
and Cloning). SalI (5' end) -NotI (3' end) cDNA was
inserted into vector PCS108
(http://mcb.berkeley.edu/labs/harland/pages/plasmids.html)
."

```

ORIGIN

```

Query Match      80.8%; Score 20.2; DB 8; Length 761;
Best Local Similarity 88.0%; Pred. No. 4.3e+02;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 1 AAAAAACACCACTGAGCTGGG 25
    |||||
Db 234 AAAAAACCACTGACTGGG 210

```

RESULT 29

```

CF219540
LOCUS           CF219540               767 bp    mRNA    linear    EST 04-AUG-2003
DEFINITION     AGENCOURT_15084592 NICHD_XGC_Emb5 Xenopus tropicalis cDNA clone
IMAGE:6989687 5', mRNA sequence.
CF219540
VERSION        CF219540.1   GI:33420248
KEYWORDS       EST.
SOURCE         Xenopus tropicalis (western clawed frog)
ORGANISM       Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
                Xenopodinae; Xenopus; Silurana.

```

REFERENCE

```

AUTHORS        NIH-MGC http://mgi.nci.nih.gov/
TITLE          National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL        Unpublished (1999)
COMMENT        Contact: Daniela S. Gerhard, Ph.D.
                Office of Cancer Genomics
                National Cancer Institute / NIH
                Bldg 31 Rm10A07 Bethesda, MD 20892
                Email: cgabbs-r@mail.nih.gov
                Tissue Procurement: Robert M. Grainger
                cDNA Library Preparation: Life Technologies, Inc.
                cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
                DNA Sequencing by: Agencourt Bioscience Corporation
                Clone distribution: MGC clone distribution information can be
                found through the I.M.A.G.E. Consortium/LLNL at:
                http://image.llnl.gov

```

```

Plate: LLAM14664 row: b column: 22
High quality sequence start: 6
High quality sequence stop: 718.
Location/Qualifiers

```

FEATURES

source

```

1..767
/organism="Xenopus tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="IMAGE:6989687"
/tissue_type="gastrula"
/dev_stage="embryo, stages 10-13"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NICHD_XGC_Emb5"
/note="Vector: pCMV-SPORT6.1; Site 1: NotI; Site 2: EcoRV;
Cloned unidirectionally. Primer: Oligo dt. Average insert
size 2.0 kb. Constructed by Invitrogen. Note: This is a
Xenopus Gene Collection (XGC) library."

```

ORIGIN

```

Query Match      80.8%; Score 20.2; DB 6; Length 767;
Best Local Similarity 88.0%; Pred. No. 4.3e+02;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 1 AAAAAACACCACTGAGCTGGG 25
    |||||
Db 733 AAAAAACCACTGACTGGG 757

```

RESULT 30

```

DN099260/c
LOCUS           DN099260               768 bp    mRNA    linear    EST 02-AUG-2005
DEFINITION     JGI CABE8370.fwd NIH_XGC_tropOval Xenopus tropicalis cDNA clone
IMAGE:7826648 5', mRNA sequence.
DN099260
VERSION        DN099260.1   GI:59769889
KEYWORDS       EST.
SOURCE         Xenopus tropicalis (western clawed frog)
ORGANISM       Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
                Xenopodinae; Xenopus; Silurana.

```

REFERENCE

```

AUTHORS        Richardson,P., Lucas,S., Rokhsar,D., Dettter,J.C., Ng,D.C.,
                Brokstein,P. and Lindquist,E.A.
TITLE          DOE Joint Genome Institute Xenopus tropicalis EST project
JOURNAL        Unpublished (2004)
COMMENT        Other ESTs: JGI CABE8370.rev
                Contact: Lindquist,E.A., Richardson,P.
                DOE Joint Genome Institute
                2800 Mitchell Drive, Walnut Creek, CA 94598, USA
                Tel: 925 296 5600
                Fax: 925 296 5710
                Email: cdna@jgi-psf.org

```

```

Tissue Procurement: Robert M. Grainger
CDNA Library Preparation: Bruce Blumberg Laboratory, University of
California, Irvine
DNA Sequencing: DOE Joint Genome Institute: http://www.jgi.doe.gov
Clone Distribution: I.M.A.G.E. Consortium/LLNL:
http://image.llnl.gov

```

Naming Conventions: EST name is generated by the concatenation of
 the JGI Clone Id and the direction of sequencing. The suffix '.fwd'
 indicates a forward sequencing read of the insert. It does not
 necessarily reflect the orientation of the insert.
 Plate: CABE 0085 row: d column: 6
 High quality sequence stop: 723.

FEATURES

source

```

1..768
/organism="Xenopus tropicalis"
/mol_type="mRNA"
/strain="N6 (Nigerian 6th generation inbred)"
/db_xref="taxon:8364"
/clone="IMAGE:7826648"
/sex="Female"
/tissue_type="Ovary"
/dev_stage="Adult"
/lab_host="ElectroMAX DH10B T1 Phage Resistant cells"
/clone_lib="NIH_XGC_tropOval"

```


/notes="Vector: pCS107; Site 1: EcoRI; Site 2: XhoI; The library was prepared from 5 ug of poly A+ RNA by oligo-dT priming (5'-ACTAGTGGCGCCCTAGCGCTGAGTCTTTTCTTTTCTTTT-3') and Stratascript reverse transcriptase. After ligation of EcoRI adapters (5'-AATTCGGCAGCAGG-3') followed by kinasing adapters and by XhoI digestion, the cDNA was size selected by chromatography on Sepharose CL-2B columns and fractions containing cDNAs larger than 1000 bp were ligated into EcoRI/XhoI-digested pCS107. Reference for library construction: Current Genomics 4, 635-644. Library constructed by Michelle Tabb and Bruce Blumberg (Dept of Developmental and Cell Biology, University of California, Irvine)."

ORIGIN

Query Match 80.8%; Score 20.2; DB 8; Length 768;
Best Local Similarity 88.0%; Pred. No. 4.3e+02;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAACACACACCTGAGCTGGG 25
||||| ||||||| ||||||| ||||||| |||||||
Db 673 AAAAAACACACACCTGACTGGG 649

RESULT 31

CX883540/c
LOCUS
DEFINITION CX883540 770 bp mRNA linear EST 04-FEB-2005
IMAGE:7673849 5', mRNA sequence.

ACCESSION CX883540
VERSION
KEYWORDS
SOURCE EST.

ORGANISM
Xenopus tropicalis (western clawed frog)

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Xenopodinae; Xenopus; Silurana.

1 (bases 1 to 770)
Richardson, P., Lucas, S., Rokhsar, D., Dettter, J.C., Ng, D.C., Brokstein, P. and Lindquist, E.A.

DOE Joint Genome Institute Xenopus tropicalis EST project
Unpublished (2004)

Contact: Lindquist, E.A., Richardson, P.
DOE Joint Genome Institute

2800 Mitchell Drive, Walnut Creek, CA 94598, USA
Tel: 925 296 5600

Fax: 925 296 5710

Email: cdna@jgi-psf.org

Tissue Procurement: Timothy Grammer (Richard M. Harland Laboratory, University of California, Berkeley:
http://tropicalis.berkeley.edu/home)

cDNA Library Preparation: DOE Joint Genome Institute:
http://www.jgi.doe.gov

DNA Sequencing: DOE Joint Genome Institute: http://www.jgi.doe.gov
Clone Distribution: I.M.A.G.E. Consortium/LLNL:
http://image.llnl.gov

Naming Conventions: EST name is generated by the concatenation of the JGI Clone Id and the direction of sequencing. The suffix '.fwd' indicates a forward sequencing read of the insert. It does not necessarily reflect the orientation of the insert.
Plate: CAAL 0233 row: e column: 15
High quality sequence stop: 761.

Location/Qualifiers
1. 770

FEATURES
source

/organism="Xenopus tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="IMAGE:7673849"
/tissue_type="Brain"
/dev_stage="Adult"
/lab_host="Electromax DH10B"
/clone_lib="NIH_XGC_tropBrn4"

/note="Vector: pCMVSPORT6; Site 1: SalI; Site 2: NotI; This library was made from dt primed cDNA and cloned into Invitrogen pCMVSPORT6 vector. The work was done at DOE Joint Genome Institute. Poly A RNA were primed with 5' GACTAGTCTAGATCGCAG CGCGCGCCCTTTTCTTTTCTTTT 3' cDNA were ligated to SalI adapter (5' TCGACCCAGCGTCCG and 5' CGACCGCTGGG), digested with NotI, size fractionated in 1.1% agarose gel electrophoresis and ligated into NotI and SalI digested pCMVSPORT6 vector."

ORIGIN

Query Match 80.8%; Score 20.2; DB 8; Length 770;
Best Local Similarity 88.0%; Pred. No. 4.3e+02;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAACACACACCTGAGCTGGG 25
||||| ||||||| ||||||| ||||||| |||||||
Db 277 AAAAAACACACACCTGACTGGG 253

RESULT 32

CX416784/c
LOCUS
DEFINITION CX416784 773 bp mRNA linear EST 06-JAN-2005
IMAGE:7578523 5', mRNA sequence.

ACCESSION CX416784.1
VERSION
KEYWORDS
SOURCE EST.

ORGANISM
Xenopus tropicalis (western clawed frog)

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Xenopodinae; Xenopus; Silurana.

1 (bases 1 to 773)
Richardson, P., Lucas, S., Rokhsar, D., Dettter, J.C., Ng, D.C., Brokstein, P. and Lindquist, E.A.

DOE Joint Genome Institute Xenopus tropicalis EST project
Unpublished (2004)

Contact: Lindquist, E.A., Richardson, P.
DOE Joint Genome Institute

2800 Mitchell Drive, Walnut Creek, CA 94598, USA
Tel: 925 296 5600

Fax: 925 296 5710

Email: cdna@jgi-psf.org

Tissue Procurement: Richard M. Harland Laboratory, University of California, Berkeley: http://tropicalis.berkeley.edu/home
cDNA Library Preparation: Richard M. Harland Laboratory, University of California, Berkeley

DNA Sequencing: DOE Joint Genome Institute: http://www.jgi.doe.gov
Clone Distribution: I.M.A.G.E. Consortium/LLNL:
http://image.llnl.gov

Naming Conventions: EST name is generated by the concatenation of the JGI Clone Id and the direction of sequencing. The suffix '.fwd' indicates a forward sequencing read of the insert. It does not necessarily reflect the orientation of the insert.
Plate: XZG 0673 row: a column: 17
High quality sequence stop: 770.

Location/Qualifiers
1. 773

FEATURES
source

/organism="Xenopus tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="IMAGE:7578523"
/tissue_type="whole embryo"
/dev_stage="Gastrula (st. 10.5-12.5)"
/lab_host="E. coli XLI-Blue derivative, Stratagene ElectroTen-Blue"
/clone_lib="NIH_XGC_tropGas7"
/note="Vector: pCS108; Site 1: SalI; Site 2: NotI; Gastrula library constructed by Russell B. Fletcher in R. Harland's lab using poly A RNA and oligo dt primers (Invitrogen SuperScript Plasmid System for cDNA Synthesis and Cloning). SalI (5' end) -NotI (3' end) cDNA was

inserted into vector pCS108
(<http://mcb.berkeley.edu/labs/harland/pages/plasmids.html>)
."

ORIGIN

Query Match 80.8%; Score 20.2; DB 8; Length 773;
Best Local Similarity 88.0%; Pred. No. 4.3e+02;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAACAACCACTGAGCTGGG 25
||||| ||||| ||||| ||||| |||||
DB 172 AAAAACAACCACTGACTGGG 148

RESULT 33

DR834532
LOCUS JGI_CABC5176.fwd NIH XGC tropFat1 linear EST 28-JUL-2005
DEFINITION IMAGE:7801607 5', mRNA sequence.

ACCESSION DR834532
VERSION DR834532.1 GI:71453472
KEYWORDS EST.
SOURCE Xenopus tropicalis (western clawed frog)
ORGANISM Xenopus tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
Xenopodinae; Xenopus; Silurana.
1 (bases 1 to 781)
Richardson,P., Lucas,S., Rokhsar,D., Dettler,J.C., Ng,D.C.,
Brokstein,P. and Lindquist,E.A.
DOE Joint Genome Institute Xenopus tropicalis EST project
Unpublished (2004)
Other_ESTS: JGI_CABC5176.rev
Contact: Lindquist,E.A., Richardson,P.
DOE Joint Genome Institute
2800 Mitchell Drive, Walnut Creek, CA 94598, USA
Tel: 925 296 5600
Fax: 925 296 5710
Email: cdna@jgi-pef.org
Tissue Procurement: Robert M. Grainger
cDNA Library Preparation: Bruce Blumberg Laboratory, University of
California, Irvine
DNA Sequencing: DOE Joint Genome Institute: <http://www.jgi.doe.gov>
Clone Distribution: I.M.A.G.E. Consortium/LLNL:
<http://image.llnl.gov>
Naming Conventions: EST name is generated by the concatenation of
the JGI Clone Id and the direction of sequencing. The suffix '.fwd'
indicates a forward sequencing read of the insert. It does not
necessarily reflect the orientation of the insert.
Plate: CABC 0053 row: P column: 21
High quality sequence stop: 701.
Location/Qualifiers
1..781
/organism="Xenopus tropicalis"
/mol_type="mRNA"
/strain="N6 (Nigerian 6th generation inbred)"
/db_xref="taxon:8364"
/clone="IMAGE:7801607"
/tissue_type="Fat body"
/dev_stage="Adult"
/lab_host="ElectroMAX DH10B T1 Phage Resistant cells"
/clone_lib="NIH XGC tropFat1"
/note="Vector: pCS107; Site 1: EcoRI; Site 2: XhoI; The
library was prepared from 5 ug of poly A+ RNA by oligo-dT
priming
(5'-ACTAGTCGGCGCTAGCGCTCGAGTTTTTTTTTTTTTTT-3') and
Stratascript reverse transcriptase. After ligation of
EcoRI adaptors (5'-AATTCGCACAGG-3') followed by kinasing
adapters and by XhoI digestion, the cDNA was size selected
by chromatography on Sepharose CL-2B columns and fractions
containing cDNAs larger than 1000 bp were ligated into
EcoRI/XhoI-digested pCS107. Reference for library
construction: Current Genomics 4, 635-644. Library

FEATURES

source

ORIGIN

Query Match 80.8%; Score 20.2; DB 8; Length 781;
Best Local Similarity 88.0%; Pred. No. 4.3e+02;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAACAACCACTGAGCTGGG 25
||||| ||||| ||||| ||||| |||||
DB 717 AAAAACAACCACTGTACTGGG 741

RESULT 34

EX729979/c
LOCUS BX729979 XGC-tadpole Xenopus tropicalis cDNA clone TTPA025dl5 5',
DEFINITION mRNA sequence.

ACCESSION BX729979
VERSION BX729979.1 GI:38402720
KEYWORDS EST.
SOURCE Xenopus tropicalis (western clawed frog)
ORGANISM Xenopus tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
Xenopodinae; Xenopus; Silurana.
1 (bases 1 to 788)
Croning,M.D.R., Ashurst,J.L., Taylor,R., Zorn,A.M. and Rogers,J.
Sanger Xenopus tropicalis EST project 2001 (11_2003)
Unpublished (2003)
Contact: Croning MDR
Sanger Institute
Hinxton, Cambridgeshire, CB10 1SA, UK
Email: trop@sanger.ac.uk
Sanger Xenopus tropicalis EST project 2001
TROPICALIS SEQUENCE ID: TTPA025dl5.plkSP6
Sequencing primer: SP6
This sequence is from a Xenopus Gene Collection (XGC) library
constructed by Nigel Garrett.
cDNA was oligo dT primed from sug of poly A+ RNA from tadpole
embryos. EcoRI-NotI cut cDNA was then ligated into pCS107 with
EcoRI at the 5' end and NotI at the 3' end.
Vector: pCS107; Site 1: EcoRI; Site 2: NotI
Host: Escherichia coli DH10B.
Location/Qualifiers
1..788
/organism="Xenopus tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="TTPA025dl5"
/dev_stage="tadpole (stage 35-40)"
/lab_host="E. coli DH10B"
/clone_lib="XGC-tadpole"
/note="Vector: pCS107; Site 1: EcoRI; Site 2: NotI; cDNA
was oligo dT primed from sug of poly A+ RNA from tadpole
embryos. EcoRI-NotI cut cDNA was then ligated into pCS107
with EcoRI at the 5' end and NotI at the 3' end"

FEATURES

source

ORIGIN

Query Match 80.8%; Score 20.2; DB 5; Length 788;
Best Local Similarity 88.0%; Pred. No. 4.3e+02;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAACAACCACTGAGCTGGG 25
||||| ||||| ||||| ||||| |||||
DB 515 AAAAACAACCACTGTACTGGG 491

RESULT 35

DN100573/c
LOCUS JGI_CABE9070.fwd NIH_XGC_tropOval linear EST 02-AUG-2005
DEFINITION

```

IMAGE:7827590 5', mRNA sequence.
DN100573
VERSION
KEYWORDS
SOURCE
ORGANISM
Xenopus tropicalis (western clawed frog)
Xenopus tropicalis
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
Xenopodinae; Xenopus; Silurana.
1 (bases 1 to 807)
Richardson,P., Lucas,S., Rokhsar,D., Dettter,J.C., Ng,D.C.,
Brokstein,P. and Lindquist,E.A.
DOE Joint Genome Institute
Unpublished (2004)
Other ESTs: JGI CABE9070.rev
Contact: Lindquist,E.A., Richardson,P.
DOE Joint Genome Institute
2800 Mitchell Drive, Walnut Creek, CA 94598, USA
Tel: 925 296 5600
Fax: 925 296 5710
Email: cdna@jgi-psf.org
Tissue Procurement: Robert M. Grainger
cDNA Library Preparation: Bruce Blumberg Laboratory, University of
California, Irvine
DNA Sequencing: DOE Joint Genome Institute; http://www.jgi.doe.gov
Clone Distribution: I.M.A.G.E. Consortium/LLNL:
http://image.llnl.gov
Naming Conventions: EST name is generated by the concatenation of
the JGI Clone Id and the direction of sequencing. The suffix '.fwd'
indicates a forward sequencing read of the insert. It does not
necessarily reflect the orientation of the insert.
Small Insert: Based upon one or more sequencing reads of this clone
where vector sequence was present at both ends, this clone has been
determined to contain a cDNA insert on the order of 600-1000 bases.
Plate: CABE 0093 row: k column: 12
High quality sequence stop: 784.
Location/Qualifiers
1. 807
/organism="Xenopus tropicalis"
/mol_type="mRNA"
/strain="N6 (Nigerian 6th generation inbred)"
/db_xref="taxon:8364"
/clone="IMAGE:7827590"
/sex="Female"
/tissue_type="Ovary"
/dev_stage="Adult"
/lab_host="ElectroMAX DH10B T1 Phage Resistant cells"
/clone_lib="NIH_XGC_tropOval"
/notes="Vector: pCS107; Site 1: EcoRI; Site 2: XhoI; The
library was prepared from 5'ug of poly A+ RNA by oligo-dT
priming
(5'-ACTAGTCGCGCCGCTAGGCTCGAGTTTCTTTTCTTTT-3') and
Stratascript reverse transcriptase. After ligation of
EcoRI adaptors (5'-AATTCGACGAGG-3') followed by kinasing
adapters and by XhoI digestion, the cDNA was size selected
by chromatography on Sepharose CL-2B columns and fractions
containing cDNAs larger than 1000 bp were ligated into
EcoRI/XhoI-digested pCS107. Reference for library
construction: Current Genomics 4, 635-644. Library
constructed by Michelle Tabb and Bruce Blumberg (Dept of
Developmental and Cell Biology, University of California,
Irvine)."
```

```

Query Match      80.8%; Score 20.2; DB 8; Length 807;
Best Local Similarity 88.0%; Pred. NO. 4.3e+02;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AAAAAACACCACTGAGCTGGG 25
||||| ||||||| |||||
Db 86 AAAAAACACCACTGAGCTGGG 62
||||| ||||||| |||||

RESULT 36
CX852874/c
LOCUS
DEFINITION
JGI CAAL9064.fwd NIH_XGC tropBrn4 Xenopus tropicalis cDNA clone
IMAGE:7666788 5', mRNA sequence.
CX852874
VERSION
KEYWORDS
SOURCE
ORGANISM
Xenopus tropicalis (western clawed frog)
Xenopus tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
Xenopodinae; Xenopus; Silurana.
1 (bases 1 to 818)
Richardson,P., Lucas,S., Rokhsar,D., Dettter,J.C., Ng,D.C.,
Brokstein,P. and Lindquist,E.A.
DOE Joint Genome Institute
Unpublished (2004)
Other ESTs: JGI CAAL9064.rev
Contact: Lindquist,E.A., Richardson,P.
DOE Joint Genome Institute
2800 Mitchell Drive, Walnut Creek, CA 94598, USA
Tel: 925 296 5600
Fax: 925 296 5710
Email: cdna@jgi-psf.org
Tissue Procurement: Timothy Grammer (Richard M. Harland Laboratory,
University of California, Berkeley:
http://tropicalis.berkeley.edu/home)
cDNA Library Preparation: DOE Joint Genome Institute:
http://www.jgi.doe.gov
DNA Sequencing: DOE Joint Genome Institute; http://www.jgi.doe.gov
Clone Distribution: I.M.A.G.E. Consortium/LLNL:
http://image.llnl.gov
Naming Conventions: EST name is generated by the concatenation of
the JGI Clone Id and the direction of sequencing. The suffix '.fwd'
indicates a forward sequencing read of the insert. It does not
necessarily reflect the orientation of the insert.
Plate: CAAL 0093 row: o column: 10
High quality sequence stop: 739.
Location/Qualifiers
1. 818
/organism="Xenopus tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="IMAGE:7666788"
/tissue_type="Brain"
/dev_stage="Adult"
/lab_host="ElectroMAX DH10B"
/clone_lib="NIH_XGC_tropBrn4"
/notes="Vector: pCMVSPORT6; Site 1: SalI; Site 2: NotI;
This library was made from dT primed cDNA and cloned into
Invitrogen pCMVSPORT6 vector. The work was done at DOE
Joint Genome Institute. Poly A RNA were primed with 5'
GACTATCTTAGATCCGAG CGCCGCCCTTTTCTTTT 3'. CDNA
were ligated to SalI adaptor (5' TCGACCCACGCGCCG and
5'CGACGCGTGGG), digested with NotI, size fractionated in
1.1% agarose gel electrophoresis and ligated into NotI and
SalI digested pCMVSPORT6 vector."
```

```

Query Match      80.8%; Score 20.2; DB 8; Length 818;
Best Local Similarity 88.0%; Pred. NO. 4.3e+02;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AAAAAACACCACTGAGCTGGG 25
||||| ||||||| |||||
Db 729 AAAAAACACCACTGAGCTGGG 705
||||| ||||||| |||||

RESULT 37
CF218992
LOCUS
DEFINITION
AGENCOURT_15085552 NICHD_XGC_Emb5 Xenopus tropicalis cDNA clone
CF218992
LOCUS
DEFINITION
```

```
IMAGE:6989266 5', mRNA sequence.
CF218992
CF218992.1 GI:33419700
EST.
Xenopus tropicalis (western clawed frog)
Xenopus tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
Xenopodinae; Xenopus; Silurana.
1 (bases 1 to 820)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Robert M. Grainger
CDNA Library Prepared by: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM14663 row: a column: 09
High quality sequence start: 6
High quality sequence stop: 735.
Location/Qualifiers
1..820
/organism="Xenopus tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="IMAGE:6989266"
/tissue_type="gastrula"
/dev_stage="embryo, stages 10-13"
/lab_host="PH10B (phage-resistant)"
/clone_lib="NICHED_XGC_Emb5"
/note="Vector: pCMV-SF0Kf6.1; Site_1: NotI; Site_2: EcoRV;
Cloned unidirectionally. Primer: Oligo dT. Average insert
size 2.0 kb. Constructed by Invitrogen. Note: This is a
Xenopus Gene Collection (XGC) library."

ORIGIN
Query Match 80.8%; Score 20.2; DB 6; Length 820;
Best Local Similarity 88.0%; Pred. No. 4.3e+02;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAAACACCACTGAGCTGGG 25
||||| ||||||| ||||||| |||||||
Db 750 AAAAAAACCACTGACTGGG 774

RESULT 38
BX758175 822 bp mRNA linear EST 10-DEC-2003
LOCUS
DEFINITION
Xenopus tropicalis (western clawed frog)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
Xenopodinae; Xenopus; Silurana.
1 (bases 1 to 822)
Croning,M.D.R., Ashurst,J.L., Taylor,P., Zorn,A.M. and Rogers,J.
Sanger Xenopus tropicalis EST project 2001 (11_2003)
Unpublished (2003)
JOURNAL
Contact: Croning MDR
Sanger Institute
Hinxton, Cambridgeshire, CB10 1SA, UK

FEATURES
source
Query Match 80.8%; Score 20.2; DB 5; Length 822;
Best Local Similarity 88.0%; Pred. No. 4.3e+02;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAAACACCACTGAGCTGGG 25
||||| ||||||| ||||||| |||||||
Db 794 AAAAAAACCACTGACTGGG 818

RESULT 39
CX403435 822 bp mRNA linear EST 06-JAN-2005
LOCUS
DEFINITION
JGI XZT50439.fwd NIH_XGC_tropad5 Xenopus tropicalis cDNA clone
IMAGE:7625675 5', mRNA sequence.
CX403435
CX403435.1 GI:57184130
EST.
Xenopus tropicalis (western clawed frog)
Xenopus tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
Xenopodinae; Xenopus; Silurana.
1 (bases 1 to 822)
Richardson,P., Lucas,S., Rokhsar,D., Detter,J.C., Ng,D.C.,
Brokstein,P. and Lindquist,E.A.
DOE Joint Genome Institute Xenopus tropicalis EST project
Unpublished (2004)
Other ESTs: JGI XZT50439.rev
Contact: Lindquist,E.A., Richardson,P.
DOE Joint Genome Institute
2800 Mitchell Drive, Walnut Creek, CA 94598, USA
Tel: 925 296 5600
Fax: 925 296 5710
Email: cdna@jgi-psf.org
Tissue Procurement: Richard M. Harland Laboratory, University of
California, Berkeley: http://tropicalis.berkeley.edu/home
CDNA Library Preparation: Richard M. Harland Laboratory, University
of California, Berkeley
DNA Sequencing: DOE Joint Genome Institute: http://www.jgi.doe.gov
Clone Distribution: I.M.A.G.E. Consortium/LLNL:
http://image.llnl.gov
Naming Conventions: EST name is generated by the concatenation of
the JGI Clone Id and the direction of sequencing. The suffix '.fwd'
indicates a forward sequencing read of the insert. It does not
necessarily reflect the orientation of the insert.
Plate: XZT 0525 row: n column: 9
```

High quality sequence stop: 748.

FEATURES

source
1..822
/organism="Xenopus tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="IMAGE:7625675"
/tissue_type="whole embryo"
/dev_stage="radpole (st. 36-41)"
/lab_host="E. coli XL1-Blue derivative, Stratagene Electropen-Blue"
/clone_lib="NIH XGC tropTad5"
/notes="vector: pCS107; Site 1: SalI; Site 2: NotI; Tadpole library constructed by Russell B. Fletcher in R. Harland's lab using poly A RNA and oligo dt primers (Invitrogen Superscript Plasmid System for cDNA Synthesis and Cloning). SalI (5' end) -NotI (3' end) cDNA was inserted into vector pCS107
(http://mcb.berkeley.edu/labs/harland/pages/plasmids.html)."

ORIGIN

Query Match 80.8%; Score 20.2; DB 8; Length 822;
Best Local Similarity 88.0%; Pred. No. 4.3e+02;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 1 AAAAAACACCACTGAGCTGGG 25
||||| ||||||| ||||||| |||||||
Db 624 AAAAAACCACTGAGCTGGG 600

RESULT 40

CR444938 830 bp mRNA linear EST 19-JUN-2004
LOCUS CR444938 XGC-tailbud Xenopus tropicalis cDNA clone TTBA065f19 3',
DEFINITION mRNA sequence.
ACCESSION CR444938.1 GI:48970525
VERSION Est.
KEYWORDS Xenopus tropicalis (western clawed frog)
SOURCE Xenopus tropicalis
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae; Xenopodinae; Xenopus; Silurana.
REFERENCE 1 (bases 1 to 830)
Croning, M.D.R., Ashurst, J.L., Taylor, R., Garrett, N. and Rogers, J.
Sanger Xenopus tropicalis EST project 2001 (2004)
TITLE Sanger Xenopus tropicalis EST project 2001 (2004)
JOURNAL Unpublished (2004)
COMMENT Contact: Croning MDR
Sanger Institute
Hinxton, Cambridgeshire, CB10 1SA, UK
Email: trop@sanger.ac.uk
Sanger Xenopus tropicalis EST project 2001
TROPICALIS_SEQUENCE_ID: TTBA065f19.q1kT7
This sequence is from a Xenopus Gene Collection (XGC) library constructed by Nigel Garrett.
Seq primer: T7.

FEATURES

source
1..830
Location/Qualifiers
/organism="Xenopus tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="TTBA065f19"
/dev_stage="tailbud (stage 28-30)"
/lab_host="Escherichia coli DH10B."
/clone_lib="XGC-tailbud"
/note="vector: pCS107; Site 1: EcoRI; Site 2: NotI; cDNA was oligo dt primed from 5ug of poly A+ RNA from tailbud. EcoRI-NotI cut cDNA was then ligated into pCS107 with EcoRI at the 5' end and NotI at the 3' end."

ORIGIN

Query Match 80.8%; Score 20.2; DB 7; Length 830;

Best Local Similarity 88.0%; Pred. No. 4.3e+02;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy

1 AAAAAACACCACTGAGCTGGG 25
||||| ||||||| ||||||| |||||||

Db

767 AAAAAACCACTGAGCTGGG 791

RESULT 41

DR863945 836 bp mRNA linear EST 29-JUL-2005
LOCUS JGI_CABG7661.fwd NIH XGC tropStol Xenopus tropicalis cDNA clone
DEFINITION IMAGE:7838710 5', mRNA sequence.
ACCESSION DR863945
VERSION DR863945.1 GI:71520535
KEYWORDS EST.
SOURCE Xenopus tropicalis (western clawed frog)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae; Xenopodinae; Xenopus; Silurana.

REFERENCE

AUTHORS Richardson, P., Lucas, S., Rokhsar, D., Dettter, J.C., Ng, D.C., Brokstein, P. and Lindquist, E.A.
TITLE DOE Joint Genome Institute Xenopus tropicalis EST project
JOURNAL Unpublished (2004)
COMMENT Other ESTs: JGI CABG7661.rev
Contact: Lindquist, E.A., Richardson, P.
DOE Joint Genome Institute
2800 Mitchell Drive, Walnut Creek, CA 94598, USA
Tel: 925 296 5600
Fax: 925 296 5710
Email: cdna@jgi-bf.org
Tissue Procurement: Robert M. Grainger
cDNA Library Preparation: Bruce Blumberg Laboratory, University of California, Irvine
DNA Sequencing: DOE Joint Genome Institute: http://www.jgi.doe.gov
Clone Distribution: I.M.A.G.E. Consortium/LLNL:
http://image.llnl.gov
Naming Conventions: EST name is generated by the concatenation of the JGI Clone Id and the direction of sequencing. The suffix '.fwd' indicates a forward sequencing read of the insert. It does not necessarily reflect the orientation of the insert.
Plate: CABG 0077 row: j column: 20
High quality sequence stop: 799.

FEATURES

source
1..836
Location/Qualifiers
/organism="Xenopus tropicalis"
/mol_type="mRNA"
/strain="N6 (Nigerian 6th generation inbred)"
/db_xref="taxon:8364"
/clone="IMAGE:7838710"
/tissue_type="Stomach"
/dev_stage="Adult"
/lab_host="NIH XGC tropStol"
/clone_lib="ElectroMAX DH10B T1 Phage Resistant cells"
/note="vector: pCS107; Site 1: EcoRI; Site 2: XhoI; The library was prepared from 5 ug of poly A+ RNA by oligo-dT priming
(5'-ACTAGTCGGCGCTAGCGCTCGAGTTTTTTTTTTTTTTTTTTT-3') and Stratascript reverse transcriptase. After ligation of EcoRI adapters (5'-AATTCGACGAGG-3') followed by kinasing adapters and by XhoI digestion, the cDNA was size selected by chromatography on Sepharose CL-2B columns and fractions containing cDNAs larger than 1000 bp were ligated into EcoRI/XhoI-digested pCS107. Reference for library construction: Current Genomics 4, 635-644. Library constructed by Michelle Tabb and Bruce Blumberg (Dept of Developmental and Cell Biology, University of California, Irvine)."

ORIGIN

Query Match 80.8%; Score 20.2; DB 8; Length 836;

Best Local Similarity 88.0%; Pred. No. 4.3e+02;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAACACCACTGAGCTGGG 25
||||| ||||| ||||| |||||
Db 777 AAAAAACCACTGACTGGG 801

RESULT 42
DR863944/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Xenopus tropicalis (western clawed frog)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
Xenopodinae; Xenopus; Silurana.
1 (bases 1 to 838)
Brookstein, P., Lucas, S., Rokhsar, D., Dettler, J.C., Ng, D.C.,
Richardson, P., and Lindquist, E.A.
Contact: Lindquist, E.A.
DOE Joint Genome Institute
2800 Mitchell Drive, Walnut Creek, CA 94598, USA
Tel: 925 236 5600
Fax: 925 236 5710
Email: cdna@jgi-psf.org
Tissue Procurement: Robert M. Grainger
CDNA Library Preparation: Bruce Blumberg Laboratory, University of
California, Irvine
DNA Sequencing: DOE Joint Genome Institute: <http://www.jgi.doe.gov>
Clone Distribution: I.M.A.G.E. Consortium/LNLN:
<http://image.lnl.gov>
Naming Conventions: EST name is generated by the concatenation of
the JGI Clone Id and the direction of sequencing. The suffix '.rev'
indicates a reverse sequencing read of the insert. It does not
necessarily reflect the orientation of the insert.
Plate: CABG 0077 row: j column: 20
High quality sequence stop: 820.

Location/Qualifiers
1. .838
/organism="Xenopus tropicalis"
/mol_type="mRNA"
/strain="N6 (Nigerian 6th generation inbred)"
/db_xref="taxon:8364"
/clone="IMAGE:7838710"
/tissue_type="Stomach"
/dev_stage="Adult"
/lab_host="ElectroMAX DH10B T1 Phage Resistant cells"
/clone_lib="NIH_XGC_tropStol"
/notes="Vector: PCS107; Site 1: EcoRI; Site 2: XhoI; The
library was prepared from 5 ug of poly A+ RNA by oligo-dT
priming
(5'-ACTAGTCGGCGCTAGCCTCGAGTCTTTTCTTTTCTTTTCTTTT-3') and
Stratascript reverse transcriptase. After ligation of
EcoRI adapters (5'-AATTCGACGAGG-3') followed by kinaasing
adapters and by XhoI digestion, the cDNA was size selected
by chromatography on Sepharose CL-2B columns and fractions
containing cDNAs larger than 1000 bp were ligated into
EcoRI/XhoI-digested PCS107. Reference for library
construction: Current Genomics 4, 635-644. Library
constructed by Michelle Rabb and Bruce Blumberg (Dept of
Developmental and Cell Biology, University of California,
Irvine)."

Query Match 80.8%; Score 20.2; DB 8; Length 838;

ORIGIN
source

Best Local Similarity 88.0%; Pred. No. 4.3e+02;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAACACCACTGAGCTGGG 25
||||| ||||| ||||| |||||
Db 236 AAAAAACCACTGACTGGG 212

RESULT 43
BX782362
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Xenopus tropicalis (western clawed frog)
Xenopus tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
Xenopodinae; Xenopus; Silurana.
1 (bases 1 to 844)
Croning, M.D.R., Ashurst, J.L., Taylor, R., Zorn, A.M. and Rogers, J.
Sanger Xenopus tropicalis EST project 2001 (11_2003)
Unpublished (2003)
Contact: Croning MDR
Sanger Institute
Hinxton, Cambridgeshire, CB10 1SA, UK
Email: trop@sanger.ac.uk
Sanger Xenopus tropicalis EST project 2001
TROPICALIS SEQUENCE ID: T8gg069p15.q1kT7
Sequencing primer: T7
This sequence is from a Xenopus Gene Collection (XGC) library
constructed by Aaron M. Zorn.
cDNA was oligo dT primed from 5ug of poly A+ RNA from egg.
EcoRI-NotI cut cDNA was then ligated into PCS107 with EcoRI at the
5' end and NotI at the 3' end.
Vector: PCS107; Site 1: EcoRI; Site 2: NotI
Host: Escherichia coli XLI-blue.

Location/Qualifiers
1. .844
/organism="Xenopus tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="T8gg069p15"
/dev_stage="egg"
/lab_host="Escherichia coli XLI-blue"
/clone_lib="XGC-egg"
/note="Vector: PCS107; Site 1: EcoRI; Site 2: NotI; cDNA
was oligo dT primed from 5ug of poly A+ RNA from egg.
EcoRI-NotI cut cDNA was then ligated into PCS107 with
EcoRI at the 5' end and NotI at the 3' end"

FEATURES
source

ORIGIN
Query Match 80.8%; Score 20.2; DB 5; Length 844;
Best Local Similarity 88.0%; Pred. No. 4.3e+02;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAACACCACTGAGCTGGG 25
||||| ||||| ||||| |||||
Db 781 AAAAAACCACTGACTGGG 805

RESULT 44
CR423927
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Xenopus tropicalis (western clawed frog)
mRNA sequence.
CR423927.1 GI:48917336
EST.
Xenopus tropicalis

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae; Xenopodinae; Xenopus; Silurana.

1 (bases 1 to 848)

Croning,M.D.R., Ashurst,J.L., Taylor,R., Garrett,N. and Rogers,J. Sanger Xenopus tropicalis EST project 2001 (2004)

Unpublished (2004)

Contact: Croning MDR

Sanger Institute

Hinxton, Cambridgeshire, CB10 1SA, UK

Email: trop@sanger.ac.uk

Sanger Xenopus tropicalis EST project 2001

TROPICALIS_SEQUENCE_ID: TTBA013109.q1kT7

This sequence is from a Xenopus Gene Collection (XGC) library

constructed by Nigel Garrett.

Seq primer: T7.

Location/Qualifiers

1..848

/organism="Xenopus tropicalis"

/mol_type="mRNA"

/db_xref="taxon:8364"

/clone="TTBA013109"

/dev_stage="tailbud (stage 28-30)"

/lab_host="Escherichia coli DH10B."

/clone_lib="XGC-tailbud"

/notes="Vector: pCS107; Site 1: EcoRI; Site 2: NotI; cDNA

was oligo dT primed from Sug of poly A+ RNA from tailbud.

EcoRI-NotI cut cDNA was then ligated into pCS107 with

EcoRI at the 5' end and NotI at the 3' end."

ORIGIN

Query Match 80.8%; Score 20.2; DB 7; Length 848;
Best Local Similarity 88.0%; Pred. No. 4.3e+02;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAACACCACTGAGCTGGG 25

||||| ||||| ||||| ||||| |||||

Db 788 AAAAAACCACTGACTGGG 812

RESULT 45

CR441514

LOCUS

DEFINITION CR441514 XGC-tailbud Xenopus tropicalis cDNA clone TTBA076108 3',

mRNA sequence.

CR441514

VERSION

KEYWORDS

SOURCE

ORGANISM

Xenopus tropicalis (western clawed frog)

Xenopus tropicalis

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;

Xenopodinae; Xenopus; Silurana.

1 (bases 1 to 851)

Croning,M.D.R., Ashurst,J.L., Taylor,R., Garrett,N. and Rogers,J.

Sanger Xenopus tropicalis EST project 2001 (2004)

Unpublished (2004)

Contact: Croning MDR

Sanger Institute

Hinxton, Cambridgeshire, CB10 1SA, UK

Email: trop@sanger.ac.uk

Sanger Xenopus tropicalis EST project 2001

TROPICALIS_SEQUENCE_ID: TTBA076108.q1kT7

This sequence is from a Xenopus Gene Collection (XGC) library

constructed by Nigel Garrett.

Seq primer: T7.

Location/Qualifiers

1..851

/organism="Xenopus tropicalis"

/mol_type="mRNA"

/db_xref="taxon:8364"

/clone="TTBA076108"

/dev_stage="tailbud (stage 28-30)"

/lab_host="Escherichia coli DH10B."

/clone_lib="XGC-tailbud"

/note="Vector: pCS107; Site 1: EcoRI; Site 2: NotI; cDNA

was oligo dT primed from Sug of poly A+ RNA from tailbud.

EcoRI-NotI cut cDNA was then ligated into pCS107 with

EcoRI at the 5' end and NotI at the 3' end."

ORIGIN

Query Match 80.8%; Score 20.2; DB 7; Length 851;
Best Local Similarity 88.0%; Pred. No. 4.3e+02;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAACACCACTGAGCTGGG 25

||||| ||||| ||||| ||||| |||||

Db 788 AAAAAACCACTGACTGGG 812

RESULT 46

BX771898

LOCUS

DEFINITION BX771898 XGC-egg Xenopus tropicalis cDNA clone Tegg044013 3', mRNA

sequence.

AX771898

VERSION

KEYWORDS

SOURCE

ORGANISM

Xenopus tropicalis (western clawed frog)

Xenopus tropicalis

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;

Xenopodinae; Xenopus; Silurana.

1 (bases 1 to 862)

Croning,M.D.R., Ashurst,J.L., Taylor,R., Zorn,A.M. and Rogers,J.

Sanger Xenopus tropicalis EST project 2001 (11_2003)

Unpublished (2003)

Contact: Croning MDR

Sanger Institute

Hinxton, Cambridgeshire, CB10 1SA, UK

Email: trop@sanger.ac.uk

Sanger Xenopus tropicalis EST project 2001

TROPICALIS_SEQUENCE_ID: Tegg044013.q1kT7

This sequence is from a Xenopus Gene Collection (XGC) library

constructed by Aaron M. Zorn.

cDNA was oligo dT primed from Sug of poly A+ RNA from egg.

EcoRI-NotI cut cDNA was then ligated into pCS107 with EcoRI at the

5' end and NotI at the 3' end.

Vector: pCS107; Site 1: EcoRI; Site 2: NotI

Host: Escherichia coli XL1-blue.

Location/Qualifiers

1..862

/organism="Xenopus tropicalis"

/mol_type="mRNA"

/db_xref="taxon:8364"

/clone="Tegg044013"

/dev_stage="egg"

/lab_host="Escherichia coli XL1-blue"

/clone_lib="XGC-egg"

/note="Vector: pCS107; Site 1: EcoRI; Site 2: NotI; cDNA

was oligo dT primed from Sug of poly A+ RNA from egg.

EcoRI-NotI cut cDNA was then ligated into pCS107 with

EcoRI at the 5' end and NotI at the 3' end"

ORIGIN

Query Match 80.8%; Score 20.2; DB 5; Length 862;
Best Local Similarity 88.0%; Pred. No. 4.3e+02;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAACACCACTGAGCTGGG 25

||||| ||||| ||||| ||||| |||||

Db 782 AAAAAACCACTGACTGGG 806

RESULT 47

```
CR568362      CR568362      862 bp      mRNA      linear      EST 19-JUL-2004
LOCUS          CR568362 XGC-tailbud-head Xenopus tropicalis cDNA clone THda017b03
DEFINITION     CR568362 3', mRNA sequence.
ACCESSION      CR568362
VERSION        CR568362.1 GI:50398439
KEYWORDS       EST.
SOURCE         Xenopus tropicalis (western clawed frog)
ORGANISM       Xenopus tropicalis
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
               Xenopodinae; Xenopus; Silurana.
REFERENCE      1 (bases 1 to 862)
AUTHORS        Croning,M.D.R., Ashurst,J.L., Taylor,R., Garrett,N. and Rogers,J.
TITLE          Sanger Xenopus tropicalis EST project 2001 (2004)
JOURNAL        Unpublished (2004)
COMMENT        Contact: Croning MDR
               Sanger Institute
               Hinxton, Cambridgeshire, CB10 1SA, UK
               Email: trop@sanger.ac.uk
               Sanger Xenopus tropicalis EST project 2001
               TROPICALIS SEQUENCE ID: THda017b03.q1kat7
               This sequence is from a Xenopus Gene Collection (XGC) library
               constructed by Nigel Garrett.
               Seq primer: T7
               Location/Qualifiers
               1..862
               /organism="Xenopus tropicalis"
               /mol_type="mRNA"
               /db_xref="taxon:8364"
               /clone="THda017b03"
               /dev_stage="tailbud head (stage 28-30)"
               /lab_host="Escherichia coli DH10B."
               /clone_lib="XGC-tailbud-head"
               /note="Vector: pCS107; Site 1: EcoRI; Site 2: NotI; cDNA
               was oligo dt primed from sug of poly A+ RNA from tailbud
               head. EcoRI-NotI cut cDNA was then ligated into pCS107
               with EcoRI at the 5' end and NotI at the 3' end."

ORIGIN
Query Match      80.8%; Score 20.2; DB 7; Length 862;
Best Local Similarity 88.0%; Pred. No. 4.3e+02;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAACACCACTGAGCTGGG 25
    ||||| ||||| ||||| |||||
Db 793 AAAAAACCACTGACTGGG 817

RESULT 48
BX755834
LOCUS          BX755834 XGC-gastrula Xenopus tropicalis cDNA clone TGas116g11 3',
DEFINITION     BX755834 mRNA sequence.
ACCESSION      BX755834
VERSION        BX755834.1 GI:39663042
KEYWORDS       EST.
SOURCE         Xenopus tropicalis (western clawed frog)
ORGANISM       Xenopus tropicalis
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
               Xenopodinae; Xenopus; Silurana.
REFERENCE      1 (bases 1 to 864)
AUTHORS        Croning,M.D.R., Ashurst,J.L., Taylor,R., Zorn,A.M. and Rogers,J.
TITLE          Sanger Xenopus tropicalis EST project 2001 (11_2003)
JOURNAL        Unpublished (2003)
COMMENT        Contact: Croning MDR
               Sanger Institute
               Hinxton, Cambridgeshire, CB10 1SA, UK
               Email: trop@sanger.ac.uk
               Sanger Xenopus tropicalis EST project 2001
               TROPICALIS SEQUENCE ID: TGas116g11.q1kt7
               Sequencing primer: T7
```

```
This sequence is from a Xenopus Gene Collection (XGC) library
constructed by Aaron M. Zorn.
cDNA was oligo dt primed from Sug of poly A+ RNA from stages 10-13
gastrulae. EcoRI-NotI cut cDNA was then ligated into pCS107 with
EcoRI at the 5' end and NotI at the 3' end.
Vector: pCS107; Site 1: EcoRI; Site 2: NotI
Host: Escherichia coli XLI-blue.

FEATURES
source
Location/Qualifiers
1..864
/organism="Xenopus tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="TGas116g11"
/dev_stage="gastrula (stages 10.5-12 mixed)"
/lab_host="Escherichia coli XLI-blue"
/clone_lib="XGC-gastrula"
/note="Vector: pCS107; Site 1: EcoRI; Site 2: NotI; cDNA
was oligo dt primed from sug of poly A+ RNA from stages
10-13 gastrulae. EcoRI-NotI cut cDNA was then ligated
into pCS107 with EcoRI at the 5' end and NotI at the 3'
end."

ORIGIN
Query Match      80.8%; Score 20.2; DB 5; Length 864;
Best Local Similarity 88.0%; Pred. No. 4.3e+02;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAACACCACTGAGCTGGG 25
    ||||| ||||| ||||| |||||
Db 801 AAAAAACCACTGACTGGG 825

RESULT 49
BX750203
LOCUS          BX750203 XGC-gastrula Xenopus tropicalis cDNA clone TGas086a19 3',
DEFINITION     BX750203 mRNA sequence.
ACCESSION      BX750203
VERSION        BX750203.1 GI:39640271
KEYWORDS       EST.
SOURCE         Xenopus tropicalis (western clawed frog)
ORGANISM       Xenopus tropicalis
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
               Xenopodinae; Xenopus; Silurana.
REFERENCE      1 (bases 1 to 869)
AUTHORS        Croning,M.D.R., Ashurst,J.L., Taylor,R., Zorn,A.M. and Rogers,J.
TITLE          Sanger Xenopus tropicalis EST project 2001 (11_2003)
JOURNAL        Unpublished (2003)
COMMENT        Contact: Croning MDR
               Sanger Institute
               Hinxton, Cambridgeshire, CB10 1SA, UK
               Email: trop@sanger.ac.uk
               Sanger Xenopus tropicalis EST project 2001
               TROPICALIS SEQUENCE ID: TGas086a19.q1kt7
               Sequencing primer: T7
               This sequence is from a Xenopus Gene Collection (XGC) library
               constructed by Aaron M. Zorn.
               cDNA was oligo dt primed from sug of poly A+ RNA from stages 10-13
               gastrulae. EcoRI-NotI cut cDNA was then ligated into pCS107 with
               EcoRI at the 5' end and NotI at the 3' end.
               Vector: pCS107; Site 1: EcoRI; Site 2: NotI
               Host: Escherichia coli XLI-blue.
               Location/Qualifiers
               1..869
               /organism="Xenopus tropicalis"
               /mol_type="mRNA"
               /db_xref="taxon:8364"
               /clone="TGas086a19"
               /dev_stage="gastrula (stages 10.5-12 mixed)"
               /lab_host="Escherichia coli XLI-blue"
               /clone_lib="XGC-gastrula"
               /note="Vector: pCS107; Site 1: EcoRI; Site 2: NotI; cDNA
```


Mon Feb 6 12:23:17 2006

was oligo dT primed from Sug of poly A+ RNA from stages 10-13 gastrulae. EcoRI-NotI cut cDNA was then ligated into pCS107 with EcoRI at the 5' end and NotI at the 3' end."

ORIGIN

Query Match 80.8%; Score 20.2; DB 5; Length 869;
Best Local Similarity 88.0%; Pred. No. 4.3e+02;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AAAAAACACCACTGAGCTGGG 25
||||| ||||| ||||| |||||
Db 788 AAAAAACCACTGACTGGG 812

RESULT 50
CR568361/c
LOCUS
DEFINITION CR568361 XGC-tailbud-head Xenopus tropicalis cDNA clone ThdA017b03
5' mRNA sequence.
CR568361
VERSION CR568361.1 GI:50398438
KEYWORDS
SOURCE
ORGANISM
Xenopus tropicalis (western clawed frog)
Xenopus tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
Xenopodinae; Xenopus; Silurana.
REFERENCE
1 (bases 1 to 872)
Croning, M.D.R., Ashurst, J.L., Taylor, R., Garrett, N. and Rogers, J.
TITLE Sahger Xenopus tropicalis EST project 2001 (2004)
JOURNAL Unpublished (2004)
COMMENT
Contact: Croning MDR
Sanger Institute
Hinxton, Cambridgeshire, CB10 1SA, UK
Email: trop@sanger.ac.uk
Sahger Xenopus tropicalis EST project 2001
TROPICALIS_SEQUENCE_ID: ThdA017b03.pikasP6
This sequence is from a Xenopus Gene Collection (XGC) library
constructed by Nigel Garrett.
Seq primer: SP6.
FEATURES
source
1..872
Location/Qualifiers
/organism="Xenopus tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="ThdA017b03"
/dev_stage="tailbud head (stage 28-30)"
/lab_host="Escherichia coli DH10B."
/clone_lib="XGC-tailbud-head"
/note="vector: pCS107; Site 1: EcoRI; Site 2: NotI; cDNA
was oligo dT primed from Sug of poly A+ RNA from tailbud
head. EcoRI-NotI cut cDNA was then ligated into pCS107
with EcoRI at the 5' end and NotI at the 3' end."

ORIGIN

Query Match 80.8%; Score 20.2; DB 7; Length 872;
Best Local Similarity 88.0%; Pred. No. 4.3e+02;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AAAAAACACCACTGAGCTGGG 25
||||| ||||| ||||| |||||
Db 713 AAAAAACCACTGACTGGG 689

Search completed: February 3, 2006, 22:01:45
Job time : 2961.67 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
 Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 3, 2006, 11:35:31 ; Search time 80.6667 Seconds
 (without alignments)
 550.897 Million cell updates/sec

Title: US-10-719-900-3
 Perfect score: 25
 Sequence: 1 aaaaacaccaccactgagctggg 25

Scoring table: IDENTITY_NUC
 Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 150 summaries

Database :

1: /cgn2_6/ptodata/1/ina/1 COMB.seq.*
 2: /cgn2_6/ptodata/1/ina/5 COMB.seq.*
 3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
 4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
 5: /cgn2_6/ptodata/1/ina/H COMB.seq.*
 6: /cgn2_6/ptodata/1/ina/PTUS COMB.seq.*
 7: /cgn2_6/ptodata/1/ina/PP COMB.seq.*
 8: /cgn2_6/ptodata/1/ina/RE COMB.seq.*
 9: /cgn2_6/ptodata/1/ina/backfile1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	18.6	74.4	6569	3	US-09-949-016-14892
2	18.6	74.4	6758	3	US-09-949-016-16871
3	18.6	74.4	15593	3	US-09-949-016-17177
4	18.6	74.4	26843	3	US-09-949-016-17208
5	18.6	74.4	106924	3	US-09-949-016-13834
6	18.2	72.8	407	3	US-09-513-999C-26191
7	18.2	72.8	601	3	US-09-949-016-69126
8	18.2	72.8	601	3	US-09-949-016-70677
9	18.2	72.8	601	3	US-09-949-016-70678
10	18.2	72.8	601	3	US-09-949-016-159778
11	18.2	72.8	601	3	US-09-949-016-195472
12	18.2	72.8	601	3	US-09-949-016-195473
13	18.2	72.8	54246	3	US-09-949-016-16206
14	18.2	72.8	58844	3	US-09-949-016-13769
15	18.2	72.8	121114	3	US-09-949-016-17232
16	18.2	72.8	154605	3	US-09-949-016-11894
17	18.2	72.8	455726	3	US-09-949-016-14157
18	18.2	72.8	481115	3	US-09-949-016-11940
19	17.8	71.2	1448	3	US-08-858-207A-71
20	17.8	71.2	1509	3	US-08-489-039A-5144
21	17.8	71.2	1542	2	US-08-865-311-1
22	17.8	71.2	1542	3	US-09-315-720-1
23	17.8	71.2	1545	3	US-09-583-110-1228
24	17.8	71.2	1566	3	US-09-107-433-1292

71.2	17.8	71.2	17702	3	US-08-961-527-7	Sequence 7, Appli
71.2	17.8	71.2	73818	3	US-09-949-016-16822	Sequence 16822, A
71.2	17.6	70.4	601	3	US-09-949-016-24377	Sequence 24377, A
70.4	17.6	70.4	601	3	US-09-949-016-134073	Sequence 134073, A
70.4	17.6	70.4	601	3	US-09-949-016-199548	Sequence 199548, A
70.4	17.6	70.4	93532	3	US-09-949-016-15944	Sequence 15944, A
70.4	17.6	70.4	221958	3	US-09-949-016-12173	Sequence 12173, A
70.4	17.6	70.4	221966	3	US-09-949-016-15498	Sequence 15498, A
70.4	17.6	70.4	373182	3	US-09-949-016-17371	Sequence 17371, A
70.4	17.6	70.4	373694	3	US-09-949-016-12062	Sequence 12062, A
69.6	17.4	69.6	705	3	US-09-248-796A-3918	Sequence 3918, Ap
69.6	17.4	69.6	882	3	US-09-248-796A-1118	Sequence 1118, Ap
69.6	17.4	69.6	22581	3	US-09-949-016-15984	Sequence 15984, A
69.6	17.4	69.6	54576	3	US-09-949-016-15954	Sequence 15954, A
69.6	17.4	69.6	54576	3	US-09-949-016-15955	Sequence 15955, A
69.6	17.4	69.6	131254	3	US-09-949-016-13734	Sequence 13734, A
68.8	17.2	68.8	601	3	US-09-949-016-22643	Sequence 22643, A
68.8	17.2	68.8	601	3	US-09-949-016-36627	Sequence 36627, A
68.8	17.2	68.8	601	3	US-09-949-016-44701	Sequence 44701, A
68.8	17.2	68.8	601	3	US-09-949-016-59992	Sequence 59992, A
68.8	17.2	68.8	601	3	US-09-949-016-172749	Sequence 172749, A
68.8	17.2	68.8	601	3	US-09-949-016-184440	Sequence 184440, A
68.8	17.2	68.8	862	3	US-09-533-559-2205	Sequence 2205, Ap
68.8	17.2	68.8	1543	2	US-08-991-948A-4	Sequence 4, Appli
68.8	17.2	68.8	2103	3	US-09-949-016-953	Sequence 953, App
68.8	17.2	68.8	2103	3	US-09-949-016-1278	Sequence 1278, Ap
68.8	17.2	68.8	7659	3	US-09-949-016-12695	Sequence 12695, A
68.8	17.2	68.8	7665	3	US-09-949-016-13020	Sequence 13020, A
68.8	17.2	68.8	30973	3	US-09-949-016-11971	Sequence 11971, A
68.8	17.2	68.8	30974	3	US-09-949-016-11571	Sequence 11571, A
68.8	17.2	68.8	30974	3	US-09-949-016-16628	Sequence 16628, A
68.8	17.2	68.8	60110	3	US-09-949-016-17338	Sequence 17338, A
68.8	17.2	68.8	60110	3	US-09-949-016-17339	Sequence 17339, A
68.8	17.2	68.8	61083	3	US-09-949-016-14144	Sequence 14144, A
68.8	17.2	68.8	61083	3	US-09-949-016-14145	Sequence 14145, A
68.8	17.2	68.8	153642	3	US-09-949-016-12174	Sequence 12174, A
68.8	17.2	68.8	153643	3	US-09-949-016-15635	Sequence 15635, A
68.8	17.2	68.8	190078	3	US-09-949-016-12707	Sequence 12707, A
68.8	17.2	68.8	190078	3	US-09-949-016-17026	Sequence 17026, A
68.8	17.2	68.8	194889	3	US-09-949-016-15654	Sequence 15654, A
68.8	17.2	68.8	343352	3	US-09-949-016-13498	Sequence 13498, A
68.8	17.2	68.8	580073	3	US-08-545-528D-1	Sequence 1, Appli
68.0	17.2	68.0	208	3	US-09-513-999C-21681	Sequence 21681, A
68.0	17.2	68.0	601	3	US-09-949-016-54261	Sequence 54261, A
68.0	17.2	68.0	601	3	US-09-949-016-54262	Sequence 54262, A
68.0	17.2	68.0	601	3	US-09-949-016-54325	Sequence 54325, A
68.0	17.2	68.0	601	3	US-09-949-016-54326	Sequence 54326, A
68.0	17.2	68.0	601	3	US-09-949-016-87785	Sequence 87785, A
68.0	17.2	68.0	601	3	US-09-949-016-160657	Sequence 160657, A
68.0	17.2	68.0	601	3	US-09-949-016-160658	Sequence 160658, A
68.0	17.2	68.0	601	3	US-09-949-016-183145	Sequence 183145, A
68.0	17.2	68.0	601	3	US-09-949-016-183886	Sequence 183886, A
68.0	17.2	68.0	639	3	US-09-533-559-5163	Sequence 5163, Ap
68.0	17.2	68.0	1558	3	US-10-104-047-1359	Sequence 1359, Ap
68.0	17.2	68.0	1747	3	US-09-227-595-25	Sequence 25, Appl
68.0	17.2	68.0	1747	3	US-09-227-595-27	Sequence 27, Appl
68.0	17.2	68.0	1747	3	US-08-595-590B-25	Sequence 25, Appl
68.0	17.2	68.0	1747	3	US-08-595-590B-27	Sequence 27, Appl
68.0	17.2	68.0	2029	2	US-07-916-098A-43	Sequence 43, Appl
68.0	17.2	68.0	2399	2	US-08-070-116A-1	Sequence 1, Appli
68.0	17.2	68.0	2399	2	US-08-557-050-1	Sequence 1, Appli
68.0	17.2	68.0	2560	2	US-07-916-098A-44	Sequence 44, Appl
68.0	17.2	68.0	9931	3	US-09-949-016-14377	Sequence 14377, A
68.0	17.2	68.0	10844	3	US-08-444-644-41	Sequence 41, Appl
68.0	17.2	68.0	10844	3	US-08-232-246A-41	Sequence 41, Appl
68.0	17.2	68.0	11194	3	US-09-949-016-14378	Sequence 14378, A
68.0	17.2	68.0	12787	3	US-09-949-016-16359	Sequence 16359, A
68.0	17.2	68.0	14189	3	US-09-949-016-15783	Sequence 15783, A
68.0	17.2	68.0	18900	3	US-09-949-016-16989	Sequence 16989, A
68.0	17.2	68.0	47677	3	US-09-949-002-668	Sequence 668, App
68.0	17.2	68.0	54986	3	US-09-949-016-16716	Sequence 16716, A
68.0	17.2	68.0	75395	3	US-09-884-890-3	Sequence 3, Appli
68.0	17.2	68.0	75395	3	US-10-274-194-3	Sequence 3, Appli
68.0	17.2	68.0	75395	3	US-10-760-407-3	Sequence 3, Appli

```

98      17 68.0 78530 3 US-09-949-016-16790
c 99      17 68.0 95561 3 US-09-949-016-12768
c 100     17 68.0 95561 3 US-09-949-016-13306
c 101     17 68.0 95561 3 US-09-949-016-13307
c 102     17 68.0 102304 3 US-09-949-016-12589
c 103     17 68.0 107820 3 US-09-792-616-1
c 104     17 68.0 110243 3 US-09-949-016-13698
c 105     17 68.0 111424 3 US-09-949-016-17014
c 106     17 68.0 112874 3 US-09-949-016-13180
c 107     17 68.0 151088 3 US-09-949-016-16240
c 108     17 68.0 190078 3 US-09-949-016-12707
c 109     17 68.0 190078 3 US-09-949-016-17026
c 110     17 68.0 236474 3 US-09-949-016-13418
c 111     17 68.0 421118 3 US-09-949-016-16297
c 112     17 68.0 670689 3 US-09-949-016-12505
c 113     17 68.0 670689 3 US-09-949-016-14207
c 114     16.8 67.2 178 3 US-09-621-976-17983
c 115     16.8 67.2 601 3 US-09-949-016-55149
c 116     16.8 67.2 601 3 US-09-949-016-62603
c 117     16.8 67.2 601 3 US-09-949-016-62604
c 118     16.8 67.2 601 3 US-09-949-016-165370
c 119     16.8 67.2 601 3 US-09-949-016-179675
c 120     16.8 67.2 601 3 US-09-949-016-179676
c 121     16.8 67.2 669 4 US-09-605-7038-2885
c 122     16.8 67.2 790 3 US-10-170-098A-1
c 123     16.8 67.2 7822 3 US-09-949-016-13339
c 124     16.8 67.2 16914 3 US-09-949-016-15972
c 125     16.8 67.2 48940 3 US-09-949-016-16402
c 126     16.8 67.2 58361 3 US-09-949-016-16755
c 127     16.8 67.2 58361 3 US-09-949-016-16756
c 128     16.8 67.2 59065 3 US-09-813-817-3
c 129     16.8 67.2 59065 3 US-09-978-197-3
c 130     16.8 67.2 59065 3 US-10-135-696-3
c 131     16.8 67.2 59065 3 US-10-820-230-3
c 132     16.8 67.2 71815 3 US-09-949-016-12501
c 133     16.8 67.2 163662 3 US-09-949-016-12545
c 134     16.8 67.2 163662 3 US-09-949-016-13546
c 135     16.8 67.2 390416 3 US-09-949-016-16923
c 136     16.6 66.4 348 3 US-09-023-655-632
c 137     16.6 66.4 399 3 US-09-621-976-10504
c 138     16.6 66.4 451 3 US-09-270-767-8403
c 139     16.6 66.4 451 3 US-09-270-767-23685
c 140     16.6 66.4 601 3 US-09-949-016-34548
c 141     16.6 66.4 601 3 US-09-949-016-41603
c 142     16.6 66.4 601 3 US-09-949-016-56245
c 143     16.6 66.4 601 3 US-09-949-016-56246
c 144     16.6 66.4 601 3 US-09-949-016-60495
c 145     16.6 66.4 601 3 US-09-949-016-78004
c 146     16.6 66.4 601 3 US-09-949-016-87388
c 147     16.6 66.4 601 3 US-09-949-016-147198
c 148     16.6 66.4 1129 3 US-09-799-451-28
c 149     16.6 66.4 1363 3 US-09-023-655-907
c 150     16.6 66.4 1523 3 US-09-019-095A-21
```

ALIGNMENTS

```

RESULT 1
US-09-949-016-14892
; Sequence 14892, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-20
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17177

US-09-949-016-17177/c
; Sequence 17177, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-20
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17177
```

```
; LENGTH: 15593
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17177

Query Match      74.4%; Score 18.6; DB 3; Length 15593;
Best Local Similarity 84.0%; Pred. No. 1.1e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAACACCACCTGAGCTGGG 25
   ||||| ||||| ||||| ||||| |||||
Db 4946 AAGAAAGAACTCCCTGAGCTGGG 4922

RESULT 4
US-09-949-016-17208/c
; Sequence 17208, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17208
; LENGTH: 26843
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17208

Query Match      74.4%; Score 18.6; DB 3; Length 26843;
Best Local Similarity 84.0%; Pred. No. 1.2e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAACACCACCTGAGCTGGG 25
   ||||| ||||| ||||| ||||| |||||
Db 5142 AAAAAAACAACACCTGAGGTGTG 5118

RESULT 5
US-09-949-016-13834/c
; Sequence 13834, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13834
; LENGTH: 106924
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature

; LOCATION: (1)...(106924)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13834

Query Match      74.4%; Score 18.6; DB 3; Length 106924;
Best Local Similarity 84.0%; Pred. No. 1.5e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAACACCACCTGAGCTGGG 25
   ||||| ||||| ||||| ||||| |||||
Db 4267 AAAAAAACAACACAGAGCTAGG 4243

RESULT 6
US-09-513-999C-26191/c
; Sequence 26191, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 26191
; LENGTH: 407
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 339
; OTHER INFORMATION: m=a or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 382
; OTHER INFORMATION: r=a or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 390
; OTHER INFORMATION: s=g or c
US-09-513-999C-26191

Query Match      72.8%; Score 18.2; DB 3; Length 407;
Best Local Similarity 87.0%; Pred. No. 99;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAACACCACCTGAGCTG 23
   ||||| ||||| ||||| ||||| |||||
Db 36 AAAAAACAAGACCACTGATCTG 14

RESULT 7
US-09-949-016-69126/c
; Sequence 69126, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
```

```
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 69126
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-69126

Query Match          72.8%; Score 18.2; DB 3; Length 601;
Best Local Similarity 80.0%; Pred. No. 1.1e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAAACACCACTGAGCTGG 24
    ||||| ||| ||| |||||
DB 450 AAAAAACACCACTGAGCTGG 428

RESULT 8
US-09-949-016-70677/c
; Sequence 70677, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 70677
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-70677

Query Match          72.8%; Score 18.2; DB 3; Length 601;
Best Local Similarity 80.0%; Pred. No. 1.1e+02;
Matches 20; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAACACCACTGAGCTGG 25
    ||||| ||| ||| |||||
DB 309 AAAAAACCAACACAGAGCTAGG 285

RESULT 9
US-09-949-016-70678/c
; Sequence 70678, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 70678
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-70678

Query Match          72.8%; Score 18.2; DB 3; Length 601;
Best Local Similarity 80.0%; Pred. No. 1.1e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAACACCACTGAGCTGG 25
    ||||| ||| ||| |||||
DB 379 AAAAAACACCACTGAGCTGG 357

RESULT 11
US-09-949-016-195472
; Sequence 195472, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 195472
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-195472

Query Match          72.8%; Score 18.2; DB 3; Length 601;
```

[illegible]

; Sequence 11894, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11894
; LENGTH: 154605
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-11894

Query Match 72.8%; Score 18.2; DB 3; Length 154605;
Best Local Similarity 87.0%; Pred. No. 2.4e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAAACACCACTGAGCTGG 24
|||||
Db 57930 AAAAAACACCACTGAGCTGG 57908

RESULT 17
US-09-949-016-14157/c
; Sequence 14157, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14157
; LENGTH: 455726
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(455726)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14157

Query Match 72.8%; Score 18.2; DB 3; Length 455726;
Best Local Similarity 87.0%; Pred. No. 2.8e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAACACCACTGAGCTG 23
|||||
Db 238030 AAAAAACACCACTGAGCTG 238008

RESULT 18
US-09-949-016-11940/c
; Sequence 11940, Application US/09949016

; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11940
; LENGTH: 481115
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(481115)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-11940

Query Match 72.8%; Score 18.2; DB 3; Length 481115;
Best Local Similarity 87.0%; Pred. No. 2.8e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAACACCACTGAGCTG 23
|||||
Db 215419 AAAAAACACCACTGAGCTG 215397

RESULT 19
US-08-858-207A-71/c
; Sequence 71, Application US/08858207A
; Patent No. 6348328
; GENERAL INFORMATION:
; APPLICANT: Black, Michael
; APPLICANT: Hodgson, John
; APPLICANT: Knowles, David
; APPLICANT: Nicholas, Richard
; APPLICANT: Stodola, Robert
; TITLE OF INVENTION: No. 6348328el Compounds
; NUMBER OF SEQUENCES: 552
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/858,207A
; FILING DATE: 09-MAY-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/017670
; FILING DATE: 14-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimm, Edward R.
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P50475
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090


```
;
; TELEX:
; INFORMATION FOR SEQ ID NO: 71:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 1448 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
; US-08-858-207A-71
;
; Query Match          71.2%; Score 17.8; DB 3; Length 1448;
; Best Local Similarity 90.5%; Pred. No. 1.8e+02;
; Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
;
; QY 3 AAACAACCACTGAGCTG 23
;      |||||
; DB 1225 AAACAACCACTGAGCTG 1205
;
; RESULT 20
; US-09-489-039A-5144
; Sequence 5144, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 5144
;   LENGTH: 1509
;   TYPE: DNA
;   ORGANISM: Klebsiella pneumoniae
; US-09-489-039A-5144
;
; Query Match          71.2%; Score 17.8; DB 3; Length 1509;
; Best Local Similarity 90.5%; Pred. No. 1.8e+02;
; Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
;
; QY 3 AAACAACCACTGAGCTG 23
;      |||||
; DB 379 AAACAACCACTGAGCTG 399
;
; RESULT 21
; US-08-865-311-1
; Sequence 1, Application US/08865311
; Patent No. 5919664
; GENERAL INFORMATION:
; APPLICANT: Pearson, Stewart C.
; TITLE OF INVENTION: No. 5919664el Compounds
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dechert, Price & Rhoads
; STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103-2793
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: Windows 95
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; FILING DATE:
; APPLICATION NUMBER: US/08/865,311
; CLASSIFICATION:
; PRIOR APPLICATION NUMBER:
;
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 1542 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: double
;   TOPOLOGY: linear
; US-09-315-720-1
;
; Query Match          71.2%; Score 17.8; DB 3; Length 1542;
; Best Local Similarity 90.5%; Pred. No. 1.8e+02;
; Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
;
; QY 3 AAACAACCACTGAGCTG 23
;      |||||
; DB 67 AAACAACCACTGAGCTG 87
;
; RESULT 22
; US-09-315-720-1
; Sequence 1, Application US/09315720
; Patent No. 6372487
; GENERAL INFORMATION:
; APPLICANT: Pearson, Stewart C.
; TITLE OF INVENTION: No. 6372487el Compounds
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dechert, Price & Rhoads
; STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103-2793
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: Windows 95
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; FILING DATE:
; APPLICATION NUMBER: US/09/315,720
; CLASSIFICATION:
; PRIOR APPLICATION NUMBER: 08/865,311
;
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Falk, Stephen T
; REGISTRATION NUMBER: 36,795
; REFERENCE/DOCKET NUMBER: GM10004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-994-2488
; TELEFAX: 215-994-2222
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 1542 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: double
;   TOPOLOGY: linear
; US-09-315-720-1
;
; Query Match          71.2%; Score 17.8; DB 3; Length 1542;
; Best Local Similarity 90.5%; Pred. No. 1.8e+02;
; Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
;
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Falk, Stephen T
; REGISTRATION NUMBER: 36,795
; REFERENCE/DOCKET NUMBER: GM10004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-994-2488
; TELEFAX: 215-994-2222
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 1542 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: double
;   TOPOLOGY: linear
; US-08-865-311-1
;
; Query Match          71.2%; Score 17.8; DB 2; Length 1542;
; Best Local Similarity 90.5%; Pred. No. 1.8e+02;
; Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
;
; QY 3 AAACAACCACTGAGCTG 23
;      |||||
; DB 67 AAACAACCACTGAGCTG 87
;
; RESULT 22
; US-09-315-720-1
; Sequence 1, Application US/09315720
; Patent No. 6372487
; GENERAL INFORMATION:
; APPLICANT: Pearson, Stewart C.
; TITLE OF INVENTION: No. 6372487el Compounds
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dechert, Price & Rhoads
; STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103-2793
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: Windows 95
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; FILING DATE:
; APPLICATION NUMBER: US/09/315,720
; CLASSIFICATION:
; PRIOR APPLICATION NUMBER: 08/865,311
;
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Falk, Stephen T
; REGISTRATION NUMBER: 36,795
; REFERENCE/DOCKET NUMBER: GM10004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-994-2488
; TELEFAX: 215-994-2222
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 1542 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: double
;   TOPOLOGY: linear
; US-09-315-720-1
;
; Query Match          71.2%; Score 17.8; DB 3; Length 1542;
; Best Local Similarity 90.5%; Pred. No. 1.8e+02;
; Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 3 AAAACAACCACTGAGCTG 23
      |||||
Db 67 AAAACAACCACTGAGCAG 87

RESULT 23
US-09-583-110-1228
; Sequence 1228, Application US/09583110
; Patent No. 6699703
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; FILE REFERENCE: PATH00-07A
; CURRENT APPLICATION NUMBER: US/09/583,110
; CURRENT FILING DATE: 2000-05-26
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 1228
; LENGTH: 1545
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
US-09-583-110-1228

Query Match 71.2%; Score 17.8; DB 3; Length 1545;
Best Local Similarity 90.5%; Pred. No. 1.8e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 AAAACAACCACTGAGCTG 23
      |||||
Db 67 AAAACAACCACTGAGCAG 87

RESULT 24
US-09-107-433-1292
; Sequence 1292, Application US/09107433
; Patent No. 6800744
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
; SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE FOR DIAGNOSTIC
; THERAPEUTICS
; NUMBER OF SEQUENCES: 5206
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: <Unknown>
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: <Unknown>
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,433
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/ 085131
; FILING DATE: May 12, 1998
; APPLICATION NUMBER: 60/051553
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-011
; TELECOMMUNICATION INFORMATION:
```

```
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 1292:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1566 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...1566
; SEQUENCE DESCRIPTION: SEQ ID NO: 1292:
US-09-107-433-1292

Query Match 71.2%; Score 17.8; DB 3; Length 1566;
Best Local Similarity 90.5%; Pred. No. 1.8e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 AAAACAACCACTGAGCTG 23
      |||||
Db 88 AAAACAACCACTGAGCAG 108
```

```
RESULT 25
US-08-961-527-7/c
; Sequence 7, Application US/08961527
; Patent No. 6420135
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 391
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,527
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19702 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-961-527-7

Query Match 71.2%; Score 17.8; DB 3; Length 19702;
Best Local Similarity 90.5%; Pred. No. 2.7e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

QY 3 AAAAAACACCACTGAGCTG 23
|||||
Db 2644 AAAAAACATCACTGAGCAG 2624

RESULT 26
US-09-949-016-16822/c
; Sequence 16822, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16822
; LENGTH: 73818
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(73818)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16822

Query Match 71.2%; Score 17.8; DB 3; Length 73818;
Best Local Similarity 90.5%; Pred. No. 3.3e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAACACCACTGAGC 21
|||||
Db 30080 AAAAAACAACTGAGC 30060

RESULT 27
US-09-949-016-24377/c
; Sequence 24377, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24377
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-24377

Query Match 70.4%; Score 17.6; DB 3; Length 601;
Best Local Similarity 83.3%; Pred. No. 1.9e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAACACCACTGAGCTG 24
|||||
Db 150 AAAAAACAACTGAGCTGG 127

RESULT 28
US-09-949-016-134073
; Sequence 134073, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 134073
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-134073

Query Match 70.4%; Score 17.6; DB 3; Length 601;
Best Local Similarity 83.3%; Pred. No. 1.9e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AAAAAACACCACTGAGCTGG 25
|||||
Db 358 AAAAAACACCACTGAGCTGG 381

RESULT 29
US-09-949-016-199548/c
; Sequence 199548, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 199548
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-199548

Query Match 70.4%; Score 17.6; DB 3; Length 601;
Best Local Similarity 83.3%; Pred. No. 1.9e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAACACCACTGAGCTGG 24
|||||
Db 150 AAAAAACAACTGAGCTGG 127

```
RESULT 30
US-09-949-016-15944/c
; Sequence 15944, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15944
; LENGTH: 93532
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15944

Query Match          70.4%; Score 17.6; DB 3; Length 93532;
Best Local Similarity 83.3%; Pred. No. 4.1e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAACACCACTGAGCTGG 24
Db 40655 AAAAAACACCAATCAGTGAGCTTG 40632

RESULT 31
US-09-949-016-12173
; Sequence 12173, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12173
; LENGTH: 221958
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12173

Query Match          70.4%; Score 17.6; DB 3; Length 221958;
Best Local Similarity 83.3%; Pred. No. 4.6e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AAAAAACACCACTGAGCTGG 25
Db 125285 AAAAAACACCACTGAGCTGG 125308

RESULT 32
US-09-949-016-15498
; Sequence 15498, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15498
; LENGTH: 221966
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15498

Query Match          70.4%; Score 17.6; DB 3; Length 221966;
Best Local Similarity 83.3%; Pred. No. 4.6e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AAAAAACACCACTGAGCTGG 25
Db 125285 AAAAAACACCACTGAGCTGG 125308

RESULT 33
US-09-949-016-17371/c
; Sequence 17371, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17371
; LENGTH: 373182
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(373182)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17371

Query Match          70.4%; Score 17.6; DB 3; Length 373182;
Best Local Similarity 83.3%; Pred. No. 4.9e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAACACCACTGAGCTGG 24
Db 67405 AAACAACAAAGAACTGAGCTGG 67382

RESULT 34
US-09-949-016-12062/c
; Sequence 12062, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
```

```

: GENERAL INFORMATION: J. Craig et al.
: APPLICANT: VENTER,
: TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
: TITLE OF INVENTION:
: FILE REFERENCE: CL001307
: CURRENT APPLICATION NUMBER: US/09/949,016
: CURRENT FILING DATE: 2000-04-14
: PRIOR APPLICATION NUMBER: 60/241,755
: PRIOR FILING DATE: 2000-10-20
: PRIOR APPLICATION NUMBER: 60/237,768

```

```
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15954
; LENGTH: 54576
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(54576)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15954

Query Match          69.6%; Score 17.4; DB 3; Length 54576;
Best Local Similarity 94.7%; Pred. No. 4.7e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAAAACACCACTGA 19
Db 54461 AAAAAACACCACTGA 54479

RESULT 39
US-09-949-016-15955
; Sequence 15955, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15955
; LENGTH: 54576
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(54576)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15955

Query Match          69.6%; Score 17.4; DB 3; Length 54576;
Best Local Similarity 94.7%; Pred. No. 4.7e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAAAACACCACTGA 19
Db 54461 AAAAAACACCACTGA 54479

RESULT 40
US-09-949-016-13734/c
; Sequence 13734, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13734
; LENGTH: 131254
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(131254)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13734

Query Match          69.6%; Score 17.4; DB 3; Length 131254;
Best Local Similarity 94.7%; Pred. No. 5.3e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 AAACAACCACTGAGCT 22
Db 127059 AAAGAACCACCACTGAGCT 127041

RESULT 41
US-09-949-016-22643
; Sequence 22643, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22643
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(54576)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-22643

Query Match          68.8%; Score 17.2; DB 3; Length 601;
Best Local Similarity 86.4%; Pred. No. 2.9e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAAACCACTGAGCTG 23
Db 509 AATAAACCACTGATCTG 530

RESULT 42
US-09-949-016-36627
; Sequence 36627, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
```

; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 36627
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-36627

Query Match 68.8%; Score 17.2; DB 3; Length 601;
Best Local Similarity 86.4%; Pred. No. 2.9e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 AAACAACCACTGAGCTGGG 25
|||||
Db 508 AAACGACCTCCACTGAGCTGGG 529

RESULT 43

US-09-949-016-44701
; Sequence 44701, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 44701
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-44701

Query Match 68.8%; Score 17.2; DB 3; Length 601;
Best Local Similarity 86.4%; Pred. No. 2.9e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 AAACAACCACTGAGCTGGG 25
|||||
Db 508 AAACGACCTCCACTGAGCTGGG 529

RESULT 44

US-09-949-016-59992
; Sequence 59992, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 59992
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-59992

Query Match 68.8%; Score 17.2; DB 3; Length 601;
Best Local Similarity 86.4%; Pred. No. 2.9e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAACCACTGAGCT 22
|||||
Db 293 AAAAAACCACTGAGCT 314

RESULT 45

US-09-949-016-172749
; Sequence 172749, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 172749
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-172749

Query Match 68.8%; Score 17.2; DB 3; Length 601;
Best Local Similarity 86.4%; Pred. No. 2.9e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAAACCACTGAGCTG 23
|||||
Db 509 ATAACACCACTGATCTG 530

RESULT 46

US-09-949-016-184440/c
; Sequence 184440, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 184440
; LENGTH: 601
; TYPE: DNA

```
; ORGANISM: Human
US-09-949-016-184440

Query Match      68.8%; Score 17.2; DB 3; Length 601;
Best Local Similarity 86.4%; Pred. No. 2.9e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 AAACAACCACTGAGCTGGG 25
||||| ||||| ||||| ||||| |||||
DB 71 AAACAACCACTGAGCCAGG 50

RESULT 47
US-09-533-559-2205
; Sequence 2205, Application US/09533559
; Patent No. 6902887
; GENERAL INFORMATION:
; APPLICANT: Randy M. Berka
; APPLICANT: Michael W. Rey
; APPLICANT: Jeffrey R. Shuster
; APPLICANT: Sakari Kauppinen
; APPLICANT: Ib Groth Clausen
; APPLICANT: Peter Bjørke Olsen
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; FILE REFERENCE: 5849.200-US
; CURRENT APPLICATION NUMBER: US/09/533,559
; CURRENT FILING DATE: 2000-03-22
; EARLIER APPLICATION NUMBER: 09/273,623
; EARLIER FILING DATE: 1999-03-22
; NUMBER OF SEQ ID NOS: 7860
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2205
; LENGTH: 862
; TYPE: DNA
; ORGANISM: Fusarium venenatum
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(862)
; OTHER INFORMATION: n = A,T,C or G
US-09-533-559-2205

Query Match      68.8%; Score 17.2; DB 3; Length 862;
Best Local Similarity 86.4%; Pred. No. 3.1e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAACAACCACTGAGCTG 23
||||| ||||| ||||| ||||| |||||
DB 228 AACACAACCACTGCACTG 249

RESULT 48
US-09-991-946A-4/c
; Sequence 4, Application US/08991946A
; Patent No. 5945306
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Guegler, Karl J.
; APPLICANT: Tang, Tom Y.
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: RAS PROTEINS
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESSES:
; ADDRESSES: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA: US/08/991,946A
; FILING DATE: December 16, 1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Cerrone, Michael C.
; REGISTRATION NUMBER: 39,132
; REFERENCE/DOCKET NUMBER: PF-0445 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1543 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: COLNNOT01
; CLONE: 607908
US-08-991-946A-4

Query Match      68.8%; Score 17.2; DB 2; Length 1543;
Best Local Similarity 79.2%; Pred. No. 3.4e+02;
Matches 19; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAACAACCACTGAGCTGG 24
||||| ||||| ||||| ||||| |||||
DB 1188 AAACAACCACTGCGCTGG 1165

RESULT 49
US-09-949-016-953
; Sequence 953, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 953
; LENGTH: 2103
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-953

Query Match      68.8%; Score 17.2; DB 3; Length 2103;
Best Local Similarity 86.4%; Pred. No. 3.5e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 AAACAACCACTGAGCTGG 25
||||| ||||| ||||| ||||| |||||
DB 2040 AAACGACCTCCACTGAACCTGG 2061

RESULT 50
US-09-949-016-1278
; Sequence 1278, Application US/09949016
; Patent No. 6812339
```



```

; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1278
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-1278

```

```

Query Match      68.8%; Score 17.2; DB 3; Length 2103;
Best Local Similarity 86.4%; Pred. No. 3.5e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

QY      4 AAACACCACCACTGAGCTGGG 25
      ||||| ||||| ||||| |||||
Db      2040 AAACGACCTCCACTGAACCTGGG 2061

```

```

Search completed: February 3, 2006, 16:31:57
Job time : 89.6667 secs

```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 3, 2006, 14:37:28 ; Search time 365.556 Seconds
(without alignments)
565.535 Million cell updates/sec

Title: US-10-719-900-3

Perfect score: 25

Sequence: 1 aaaaacaccaccactgagctggg 25

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database : Published Applications_NA_Main:*

1: /cgn2_6/prodata/1/pubpna/US07_PUBCOMB.seq:*

2: /cgn2_6/prodata/1/pubpna/US08_PUBCOMB.seq:*

3: /cgn2_6/prodata/1/pubpna/US09A_PUBCOMB.seq:*

4: /cgn2_6/prodata/1/pubpna/US09B_PUBCOMB.seq:*

5: /cgn2_6/prodata/1/pubpna/US10A_PUBCOMB.seq:*

6: /cgn2_6/prodata/1/pubpna/US10B_PUBCOMB.seq:*

7: /cgn2_6/prodata/1/pubpna/US10C_PUBCOMB.seq:*

8: /cgn2_6/prodata/1/pubpna/US10D_PUBCOMB.seq:*

9: /cgn2_6/prodata/1/pubpna/US10E_PUBCOMB.seq:*

10: /cgn2_6/prodata/1/pubpna/US11_PUBCOMB.seq:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	25	100.0	25	8	US-10-719-900-3
2	23.4	93.6	25	8	US-10-719-900-4
3	20.2	80.8	512	5	US-10-027-632-48258
4	20.2	80.8	512	6	US-10-027-632-48258
5	20.2	80.8	568	4	US-09-925-065A-201949
6	20.2	80.8	620	5	US-10-027-632-75571
7	20.2	80.8	620	5	US-10-027-632-76226
8	20.2	80.8	620	6	US-10-027-632-75571
9	20.2	80.8	620	6	US-10-027-632-76226
10	19.8	79.2	556	4	US-09-925-065A-148340
11	19.2	76.8	201	8	US-10-719-993-34439
12	19.2	76.8	201	8	US-10-719-993-34440
13	19.2	76.8	503	3	US-09-918-995-31484
14	19.2	76.8	546025	8	US-10-719-993-31484
15	18.8	75.2	42061	7	US-10-322-281-770
16	18.8	75.2	96898	6	US-10-417-476-3
17	18.8	75.2	367378	6	US-10-312-841-2
18	18.6	74.4	90	9	US-10-467-851-264
19	18.6	74.4	276	8	US-10-425-115-58282
20	18.6	74.4	333	7	US-10-242-535A-8589
21	18.6	74.4	333	7	US-10-085-783A-8589
22	18.6	74.4	540	7	US-10-276-774-680
23	18.6	74.4	709	4	US-09-925-065A-934362

24	18.6	74.4	755	5	US-10-027-632-129384	Sequence 129384,
25	18.6	74.4	755	7	US-10-027-632-129384	Sequence 129384,
c 26	18.6	74.4	777	6	US-10-343-650A-221	Sequence 221, App
c 27	18.6	74.4	882	5	US-10-027-633-129760	Sequence 129760,
c 28	18.6	74.4	882	5	US-10-027-633-129761	Sequence 129761,
c 29	18.6	74.4	882	5	US-10-027-633-129762	Sequence 129762,
c 30	18.6	74.4	882	5	US-10-027-633-129760	Sequence 129760,
c 31	18.6	74.4	882	6	US-10-027-633-129760	Sequence 129760,
c 32	18.6	74.4	882	6	US-10-027-633-129761	Sequence 129761,
c 33	18.6	74.4	882	6	US-10-027-633-129762	Sequence 129762,
c 34	18.6	74.4	882	6	US-10-027-633-129760	Sequence 129760,
c 35	18.6	74.4	882	6	US-10-027-633-129760	Sequence 129760,
c 36	18.6	74.4	1065	6	US-10-220-382-32	Sequence 32, Appl
c 37	18.6	74.4	1329	3	US-09-965-422-1	Sequence 1, Appl
c 38	18.6	74.4	1450	5	US-10-012-600B-103	Sequence 103, App
c 39	18.6	74.4	2028	3	US-09-981-566A-43	Sequence 43, Appl
c 40	18.6	74.4	2464	3	US-09-764-891-6728	Sequence 6728, Ap
c 41	18.6	74.4	2464	3	US-09-764-891-6729	Sequence 6729, Ap
c 42	18.6	74.4	3029	8	US-10-723-860-7237	Sequence 7237, Ap
c 43	18.6	74.4	3154	6	US-10-172-118-1308	Sequence 1308, Ap
c 44	18.6	74.4	3154	7	US-10-342-887-1308	Sequence 1308, Ap
c 45	18.6	74.4	22929	6	US-10-017-161-591	Sequence 591, App
c 46	18.6	74.4	122923	5	US-10-087-192-268	Sequence 268, App
c 47	18.6	74.4	128993	8	US-10-484-577-681	Sequence 681, App
c 48	18.6	74.4	177380	8	US-10-484-577-683	Sequence 683, App
c 49	18.4	73.6	607	5	US-10-027-632-279949	Sequence 279949,
c 50	18.4	73.6	607	5	US-10-027-632-279950	Sequence 279950,
c 51	18.4	73.6	607	6	US-10-027-632-279949	Sequence 279949,
c 52	18.4	73.6	607	6	US-10-027-632-279950	Sequence 279950,
c 53	18.4	73.6	2097	7	US-10-282-122A-40836	Sequence 40836, A
c 54	18.4	73.6	47756	7	US-10-322-696-148	Sequence 148, App
c 55	18.2	72.8	201	8	US-10-719-993-28266	Sequence 28266, A
c 56	18.2	72.8	485	7	US-10-424-599-60162	Sequence 60162, A
c 57	18.2	72.8	553	4	US-09-925-065A-832430	Sequence 832430,
c 58	18.2	72.8	592	4	US-09-925-065A-448402	Sequence 448402,
c 59	18.2	72.8	592	4	US-09-925-065A-448403	Sequence 448403,
c 60	18.2	72.8	598	3	US-09-864-761-7372	Sequence 7372, Ap
c 61	18.2	72.8	599	9	US-10-972-079-71589	Sequence 71589, A
c 62	18.2	72.8	600	9	US-10-972-079-71590	Sequence 71590, A
c 63	18.2	72.8	600	9	US-10-972-079-71591	Sequence 71591, A
c 64	18.2	72.8	600	9	US-10-972-079-71592	Sequence 71592, A
c 65	18.2	72.8	600	9	US-10-972-079-92735	Sequence 92735, A
c 66	18.2	72.8	616	4	US-09-925-065A-952141	Sequence 952141,
c 67	18.2	72.8	652	5	US-10-027-632-242984	Sequence 242984,
c 68	18.2	72.8	652	6	US-10-027-632-242984	Sequence 242984,
c 69	18.2	72.8	1165	7	US-10-424-599-106254	Sequence 106254,
c 70	18.2	72.8	1383	4	US-09-925-065A-88703	Sequence 88703, A
c 71	18.2	72.8	4624	10	US-11-097-143-30992	Sequence 30992, A
c 72	18.2	72.8	326014	3	US-09-731-231A-3	Sequence 3, Appli
c 73	18.2	72.8	326014	7	US-10-751-985-3	Sequence 3, Appli
c 74	18.2	72.8	493999	8	US-10-719-993-6787	Sequence 6787, Ap
c 75	17.8	71.2	108	3	US-09-864-761-28125	Sequence 28125, A
c 76	17.8	71.2	352	3	US-09-764-864-659	Sequence 659, App
c 77	17.8	71.2	352	3	US-09-764-891-419	Sequence 419, App
c 78	17.8	71.2	360	6	US-10-091-414-26	Sequence 26, Appl
c 79	17.8	71.2	362	3	US-09-864-761-11537	Sequence 11537, A
c 80	17.8	71.2	427	7	US-10-437-963-31341	Sequence 31341, A
c 81	17.8	71.2	435	9	US-10-450-763-3205	Sequence 3205, Ap
c 82	17.8	71.2	504	6	US-10-029-386-6909	Sequence 6909, Ap
c 83	17.8	71.2	606	6	US-10-029-386-22987	Sequence 22987, A
c 84	17.8	71.2	1485	7	US-10-282-122A-23394	Sequence 23394, A
c 85	17.8	71.2	1542	8	US-10-472-928-725	Sequence 725, App
c 86	17.8	71.2	1545	7	US-10-282-122A-37633	Sequence 37633, A
c 87	17.8	71.2	1566	9	US-10-617-320-1292	Sequence 1292, Ap
c 88	17.8	71.2	1776	9	US-10-450-763-29300	Sequence 29300, A
c 89	17.8	71.2	1895	3	US-09-764-864-230	Sequence 230, App
c 90	17.8	71.2	2056	6	US-10-172-118-1662	Sequence 1662, Ap
c 91	17.8	71.2	2056	6	US-10-342-887-1662	Sequence 1662, Ap
c 92	17.8	71.2	5664	9	US-10-450-763-3207	Sequence 3207, Ap
c 93	17.8	71.2	19702	2	US-08-961-527-7	Sequence 7, Appli
c 94	17.8	71.2	19702	7	US-10-158-844-7	Sequence 7, Appli
c 95	17.8	71.2	26006	3	US-09-764-869-1963	Sequence 1963, Ap
c 96	17.8	71.2	26006	3	US-09-764-864-1638	Sequence 1638, Ap

```
c 97 17.8 71.2 26006 3 US-09-764-891-6285 Sequence 6285, Ap
c 98 17.8 71.2 26006 5 US-10-091-504-1963 Sequence 1963, Ap
c 99 17.8 71.2 26006 6 US-10-091-414-242 Sequence 242, App
c 100 17.8 71.2 26006 6 US-10-242-355-997 Sequence 997, App
c 101 17.8 71.2 26006 3 US-10-227-577-1963 Sequence 1963, Ap
c 102 17.8 71.2 26013 3 US-09-764-869-1961 Sequence 1961, Ap
c 103 17.8 71.2 26013 3 US-09-764-864-1636 Sequence 1636, Ap
c 104 17.8 71.2 26013 3 US-09-764-891-6283 Sequence 6283, Ap
c 105 17.8 71.2 26013 5 US-10-091-504-1961 Sequence 1961, Ap
c 106 17.8 71.2 26013 6 US-10-091-414-240 Sequence 240, App
c 107 17.8 71.2 26013 6 US-10-242-355-995 Sequence 995, App
c 108 17.8 71.2 26013 6 US-10-227-577-1961 Sequence 1961, Ap
c 109 17.8 71.2 26018 3 US-09-764-869-1962 Sequence 1962, Ap
c 110 17.8 71.2 26018 3 US-09-764-864-1637 Sequence 1637, Ap
c 111 17.8 71.2 26018 3 US-09-764-891-6284 Sequence 6284, Ap
c 112 17.8 71.2 26018 5 US-10-091-504-1962 Sequence 1962, Ap
c 113 17.8 71.2 26018 6 US-10-091-414-241 Sequence 241, App
c 114 17.8 71.2 26018 6 US-10-242-355-996 Sequence 996, App
c 115 17.8 71.2 26018 6 US-10-227-577-1962 Sequence 1962, Ap
c 116 17.8 71.2 26043 7 US-10-304-113-4 Sequence 4, Appli
c 117 17.8 71.2 230101 8 US-10-719-993-6829 Sequence 6829, Ap
c 118 17.8 71.2 2162598 8 US-10-472-928-4979 Sequence 4979, Ap
c 119 17.6 70.4 201 8 US-10-719-993-36574 Sequence 36574, A
c 120 17.6 70.4 265 3 US-09-560-863-148 Sequence 148, App
c 121 17.6 70.4 501 5 US-10-027-632-107363 Sequence 107363,
c 122 17.6 70.4 522 4 US-10-027-632-107363 Sequence 107363,
c 123 17.6 70.4 522 4 US-09-925-065A-64391 Sequence 64391, A
c 124 17.6 70.4 522 4 US-09-925-065A-64392 Sequence 64392, A
c 125 17.6 70.4 538 5 US-10-027-632-199273 Sequence 199273,
c 126 17.6 70.4 538 5 US-10-027-632-199273 Sequence 199273,
c 127 17.6 70.4 600 4 US-09-925-065A-116076 Sequence 116076,
c 128 17.6 70.4 600 4 US-09-925-065A-116077 Sequence 116077,
c 129 17.6 70.4 600 9 US-10-972-079-20258 Sequence 20258, A
c 130 17.6 70.4 600 9 US-10-972-079-20259 Sequence 20259, A
c 131 17.6 70.4 600 9 US-10-972-079-41751 Sequence 41751, A
c 132 17.6 70.4 600 9 US-10-972-079-53149 Sequence 53149, A
c 133 17.6 70.4 600 9 US-10-972-079-53150 Sequence 53150, A
c 134 17.6 70.4 617 4 US-09-925-065A-387867 Sequence 387867,
c 135 17.6 70.4 617 4 US-09-925-065A-387868 Sequence 387868,
c 136 17.6 70.4 617 4 US-09-925-065A-387869 Sequence 387869,
c 137 17.6 70.4 618 5 US-10-027-632-160009 Sequence 160009,
c 138 17.6 70.4 618 5 US-10-027-632-160009 Sequence 160009,
c 139 17.6 70.4 631 4 US-09-925-065A-687630 Sequence 687630,
c 140 17.6 70.4 631 4 US-09-925-065A-687631 Sequence 687631,
c 141 17.6 70.4 636 4 US-09-925-065A-359335 Sequence 359335,
c 142 17.6 70.4 764 3 US-09-770-445-941 Sequence 941, App
c 143 17.6 70.4 782 9 US-10-779-543-3972 Sequence 3972, Ap
c 144 17.6 70.4 1323 8 US-10-474-793-657 Sequence 657, App
c 145 17.6 70.4 1337 4 US-09-925-065A-704583 Sequence 704583,
c 146 17.6 70.4 1337 4 US-09-925-065A-704584 Sequence 704584,
c 147 17.6 70.4 1337 4 US-09-925-065A-704585 Sequence 704585,
c 148 17.6 70.4 1337 4 US-09-925-065A-704586 Sequence 704586,
c 149 17.6 70.4 3399 7 US-10-424-593-113687 Sequence 113687,
c 150 17.6 70.4 7386 7 US-10-114-270-193 Sequence 193, App
```

ALIGNMENTS

```
RESULT 1
US-10-719-900-3
; Sequence 3, Application US/10719900
; Publication No. US20050026164A1
; GENERAL INFORMATION:
; APPLICANT: Xue Mei Zhou
; TITLE OF INVENTION: Methods of Genetic Analysis of Mouse
; FILE REFERENCE: 3528.1
; CURRENT APPLICATION NUMBER: US/10719,900
; CURRENT FILING DATE: 2003-11-20
; PRIOR APPLICATION NUMBER: 60/427,808
; PRIOR FILING DATE: 2002 11 20
; NUMBER OF SEQ ID NOS: 982914
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1

US-10-719-900-3
; SEQ ID NO 3
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-719-900-3
Query Match 100.0%; Score 25; DB 8; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAAAAAACCACTGAGCTGGG 25
|||||
Db 1 AAAAAAACCACTGAGCTGGG 25
|||||

RESULT 2
US-10-719-900-4
; Sequence 4, Application US/10719900
; Publication No. US20050026164A1
; GENERAL INFORMATION:
; APPLICANT: Xue Mei Zhou
; TITLE OF INVENTION: Methods of Genetic Analysis of Mouse
; FILE REFERENCE: 3528.1
; CURRENT APPLICATION NUMBER: US/10719,900
; CURRENT FILING DATE: 2003-11-20
; PRIOR APPLICATION NUMBER: 60/427,808
; PRIOR FILING DATE: 2002 11 20
; NUMBER OF SEQ ID NOS: 982914
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 4
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-719-900-4
Query Match 93.6%; Score 23.4; DB 8; Length 25;
Best Local Similarity 96.0%; Pred. No. 0.86;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 AAAAAAACCACTGAGCTGGG 25
|||||
Db 1 AAAAAAACCACTGAGCTGGG 25
|||||

RESULT 3
US-10-027-632-48258/c
; Sequence 48258, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 48258
; LENGTH: 512
```


Query Match	76.8%	Score 19.2;	DB 8;	Length 201;
Best Local Similarity	87.5%	Pred. No. 91;		
Matches 21: Conservative	0:	Mismatches	3:	Indels 0: Gaps 0:

RESULT 15

US-10-322-281-770/c
; Sequence 770, Application US/10322281
; Publication No. US20040126762A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: 529452001000
; CURRENT APPLICATION NUMBER: US/10/322,281
; CURRENT FILING DATE: 2002-12-17
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 770
; LENGTH: 42061
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(42061)
; OTHER INFORMATION: n = A, T, C or G
US-10-322-281-770

Query Match 75.2%; Score 18.8; DB 7; Length 42061;
Best Local Similarity 90.9%; Pred. No. 2.5e+02;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAACACACCACTGAGCT 22

Db 10151 AAAAAACACACCACTGAGCT 10130

RESULT 16

US-10-417-476-3
; Sequence 3, Application US/10417476
; Publication No. US20040002102A1
; GENERAL INFORMATION:
; APPLICANT: Litman, Gary W.
; APPLICANT: Hawke, No. US20040002102A11 A.
; APPLICANT: Yoder, Jeffrey A.
; TITLE OF INVENTION: B1VM (Basic, Immunoglobulin-Like Variable Motif-Containing) Gene,
; FILE REFERENCE: Transcriptional Products, and Uses Thereof
; FILE REFERENCE: USF-103X
; CURRENT APPLICATION NUMBER: US/10/417,476
; CURRENT FILING DATE: 2003-04-16
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 96898
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (466)..(702)
; OTHER INFORMATION: Inverse complement of MGC5302 Exon 6
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (823)..(997)
; OTHER INFORMATION: Inverse complement of MGC5302 Exon 5
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2904)..(2985)
; OTHER INFORMATION: Inverse complement of MGC5302 Exon 4
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (3164)..(3366)
; OTHER INFORMATION: Inverse complement of MGC5302 Exon 3
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (6395)..(6570)
; OTHER INFORMATION: Inverse complement of MGC5302 Exon 2
; FEATURE:
; NAME/KEY: misc feature

; LOCATION: (8013)..(8312)
; OTHER INFORMATION: Genomic fragment identified as part of CpG island (Genbank Z59762)
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (8059)..(8572)
; OTHER INFORMATION: Inverse complement of MGC5302 Exon 1
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (8233)..(8235)
; OTHER INFORMATION: Inverse complement of MGC5302 translation initiation codon (ATG)
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (8614)..(9086)
; OTHER INFORMATION: B1VM Exon A - untranslated
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (9019)..(9033)
; OTHER INFORMATION: B1VM Exon A alternative splice donor site
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (9077)..(9091)
; OTHER INFORMATION: B1VM Exon A splice donor site
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (14862)..(14876)
; OTHER INFORMATION: B1VM Exon B splice acceptor site
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (14872)..(14955)
; OTHER INFORMATION: B1VM Exon B - untranslated
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (14946)..(14960)
; OTHER INFORMATION: B1VM Exon B splice donor site
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (16309)..(16309)
; OTHER INFORMATION: n = a, c, g, or t.
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (16701)..(16715)
; OTHER INFORMATION: B1VM Exon 1 splice acceptor site
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (16711)..(17310)
; OTHER INFORMATION: B1VM Exon 1
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (16833)..(16835)
; OTHER INFORMATION: B1VM translation initiation codon (ATG)
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (17301)..(17315)
; OTHER INFORMATION: B1VM Exon 1 splice donor site
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (25983)..(25997)
; OTHER INFORMATION: B1VM Exon 2 splice acceptor site
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (25993)..(26119)
; OTHER INFORMATION: B1VM Exon 2
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (26110)..(26124)
; OTHER INFORMATION: B1VM Exon 2 splice donor site
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (30552)..(30606)
; OTHER INFORMATION: B1VM Exon 3 splice acceptor site
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (30602)..(30697)


```
OTHER INFORMATION: BIVM Exon 3
FEATURE:
NAME/KEY: misc feature
LOCATION: (30688)..(30702)
OTHER INFORMATION: BIVM Exon 3 splice donor site
FEATURE:
NAME/KEY: misc feature
LOCATION: (31298)..(31312)
OTHER INFORMATION: BIVM Exon 4 splice acceptor site
FEATURE:
NAME/KEY: misc feature
LOCATION: (31308)..(31412)
OTHER INFORMATION: BIVM Exon 4
FEATURE:
NAME/KEY: misc feature
LOCATION: (31403)..(31417)
OTHER INFORMATION: BIVM Exon 4 splice donor site
FEATURE:
NAME/KEY: misc feature
LOCATION: (31620)..(31634)
OTHER INFORMATION: BIVM Exon 5 splice acceptor site
FEATURE:
NAME/KEY: misc feature
LOCATION: (31630)..(31724)
OTHER INFORMATION: BIVM Exon 5
FEATURE:
NAME/KEY: misc feature
LOCATION: (31715)..(31729)
OTHER INFORMATION: BIVM Exon 5 splice donor site
FEATURE:
NAME/KEY: misc feature
LOCATION: (41120)..(41134)
OTHER INFORMATION: BIVM Exon 6 splice acceptor site
FEATURE:
NAME/KEY: misc feature
LOCATION: (41130)..(41262)
OTHER INFORMATION: BIVM Exon 6
FEATURE:
NAME/KEY: misc feature
LOCATION: (41253)..(41267)
OTHER INFORMATION: BIVM Exon 6 splice donor site
FEATURE:
NAME/KEY: misc feature
LOCATION: (44021)..(44035)
OTHER INFORMATION: BIVM Exon 7 splice acceptor site
FEATURE:
NAME/KEY: misc feature
LOCATION: (44031)..(44117)
OTHER INFORMATION: BIVM Exon 7
FEATURE:
NAME/KEY: misc feature
LOCATION: (44108)..(44122)
OTHER INFORMATION: BIVM Exon 7 splice donor site
FEATURE:
NAME/KEY: misc feature
LOCATION: (48198)..(48212)
OTHER INFORMATION: BIVM Exon 8 splice acceptor site
FEATURE:
NAME/KEY: misc feature
LOCATION: (48208)..(48304)
OTHER INFORMATION: BIVM Exon 8
FEATURE:
NAME/KEY: misc feature
LOCATION: (48295)..(48309)
OTHER INFORMATION: BIVM Exon 8 splice donor site
FEATURE:
NAME/KEY: misc feature
LOCATION: (49127)..(49141)
OTHER INFORMATION: BIVM Exon 9 splice acceptor site
FEATURE:
NAME/KEY: misc feature
LOCATION: (49137)..(51096)
OTHER INFORMATION: BIVM Exon 9
```

```
FEATURE:
NAME/KEY: misc feature
LOCATION: (49428)..(49430)
OTHER INFORMATION: BIVM translation termination codon (TGA)
FEATURE:
NAME/KEY: misc feature
LOCATION: (50039)..(50248)
OTHER INFORMATION: Alu sequence
FEATURE:
NAME/KEY: misc feature
LOCATION: (55216)..(55975)
OTHER INFORMATION: ERCC5 Exon 1
FEATURE:
NAME/KEY: misc feature
```

Query Match 75.2%; Score 18.8; DB 6; Length 96898;
Best Local Similarity 90.9%; Pred. No. 2.8e+02;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAACACCACTGAGCT 22
|||||
DB 5945 AAAAAACACCACTGAGCT 5966
|||||

RESULT 17

US-10-312-841-2/c
; Sequence 2, Application US/10312841
; Publication No. US20030186277A1
; GENERAL INFORMATION:
; APPLICANT: Epigenomics AG
; TITLE OF INVENTION: Diagnose von bedeutenden genetischen Parametern innerhalb des MHA
; FILE REFERENCE: E01/1208/WO
; CURRENT APPLICATION NUMBER: US/10/312,841
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 2
; SEQ ID NO 2
; LENGTH: 3673778
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (379615)
US-10-312-841-2

Query Match 75.2%; Score 18.8; DB 6; Length 3673778;
Best Local Similarity 90.9%; Pred. No. 3.8e+02;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAACACCACTGAGCT 22
|||||
DB 844278 AAAAAACACCACTAATCT 844257
|||||

RESULT 18

US-10-467-851-264/c
; Sequence 264, Application US/10467851
; Publication No. US20050221303A1
; GENERAL INFORMATION:
; APPLICANT: TELERMAN, Adam
; APPLICANT: AMSON, Robert
; APPLICANT: TUIJNDER, Marcel,
; APPLICANT: SUSINI, Laurent
; TITLE OF INVENTION: SEQUENCES INVOLVED IN PHENOMENA OF TUMOUR SUPPRESSION,
; TITLE OF INVENTION: TUMOUR REVERSION, APOPTOSIS AND/OR VIRUS RESISTANCE
; TITLE OF INVENTION: AND THEIR USE AS MEDICINES
; FILE REFERENCE: 11416-014-999
; CURRENT APPLICATION NUMBER: US/10/467,851
; CURRENT FILING DATE: 2003-08-13
; PRIOR APPLICATION NUMBER: PCT/FR 02/00 543
; PRIOR FILING DATE: 2002-02-13
; PRIOR APPLICATION NUMBER: FR 01/01 925

```
; PRIOR FILING DATE: 2001-02-13
; NUMBER OF SEQ ID NOS: 1020
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 264
; LENGTH: 90
; TYPE: DNA
; ORGANISM: HOMO SAPIENS
US-10-467-851-264

Query Match          74.4%; Score 18.6; DB 9; Length 90;
Best Local Similarity 84.0%; Pred. No. 1.6e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAACACCACTGAGCTGGG 25
   |||||
DB 43 AAAAAACACCACTGAGCTGGG 19
   |||||

RESULT 19
US-10-425-115-58262/c
; Sequence 58262, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 58262
; LENGTH: 276
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(276)
; OTHER INFORMATION: unsure at all n locations
; OTHER INFORMATION: Clone ID: MRT4577_153111C.1
US-10-425-115-58262

Query Match          74.4%; Score 18.6; DB 8; Length 276;
Best Local Similarity 84.0%; Pred. No. 1.8e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAACACCACTGAGCTGGG 25
   |||||
DB 129 AAAAAACAGGACCACTGAGCTGTG 105
   |||||

RESULT 20
US-10-242-535A-8589
; Sequence 8589, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liew, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2005
; CURRENT APPLICATION NUMBER: US/10/242,535A
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 10/085,783
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
```

```
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 8589
; LENGTH: 333
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (7)..(7)
; OTHER INFORMATION: n is a, c, g, or t
; NAME/KEY: misc feature
; LOCATION: (315)..(315)
; OTHER INFORMATION: n is a, c, g, or t
US-10-242-535A-8589

Query Match          74.4%; Score 18.6; DB 7; Length 333;
Best Local Similarity 84.0%; Pred. No. 1.8e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAACACCACTGAGCTGGG 25
   |||||
DB 279 AAAAAACACCACTGAGCTGGG 303
   |||||

RESULT 21
US-10-085-783A-8589
; Sequence 8589, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liew, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2002
; CURRENT APPLICATION NUMBER: US/10/085,783A
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 8589
; LENGTH: 333
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (7)..(7)
; OTHER INFORMATION: n is a, c, g, or t
; NAME/KEY: misc feature
; LOCATION: (315)..(315)
; OTHER INFORMATION: n is a, c, g, or t
US-10-085-783A-8589

Query Match          74.4%; Score 18.6; DB 7; Length 333;
Best Local Similarity 84.0%; Pred. No. 1.8e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAACACCACTGAGCTGGG 25
   |||||
DB 279 AAAAAACACCACTGAGCTGGG 303
   |||||

RESULT 22
US-10-278-774-680
; Sequence 680, Application US/10276774
; Publication No. US20040053245A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
```

```

; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 129384
; LENGTH: 755
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-129384

Query Match          74.4%; Score 18.6; DB 5; Length 755;
Best Local Similarity 84.0%; Pred. No. 2e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAACAACCACTGAGCTGGG 25
    ||||| ||||| ||||| |||||
DB 588 AAAAAACAACAAACTGAACTGGG 612

RESULT 25
US-10-027-632-129384
; Sequence 129384, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; POLYMERIZATION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 129384
; LENGTH: 755
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-129384

Query Match          74.4%; Score 18.6; DB 6; Length 755;
Best Local Similarity 84.0%; Pred. No. 2e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAACAACCACTGAGCTGGG 25
    ||||| ||||| ||||| |||||
DB 588 AAAAAACAACAAACTGAACTGGG 612

```

RESULT 26
US-10-343-650A-221/c
; Sequence 221, Application US/10343650A
; Publication No. US20040067499A1
; GENERAL INFORMATION:
; APPLICANT: HAGA, TATSUYA
; TITLE OF INVENTION: NOVEL G-PROTEIN COUPLED RECEPTOR
; FILE REFERENCE: 31671-186347
; CURRENT APPLICATION NUMBER: US/10/343,650A
; PRIOR FILING DATE: 2003-07-21
; PRIOR APPLICATION NUMBER: JP 2000/237818
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: JP 2001/34434
; PRIOR FILING DATE: 2001-02-13
; NUMBER OF SEQ ID NOS: 694
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 221
; LENGTH: 777
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(777)
US-10-343-650A-221

Query Match 74.4%; Score 18.6; DB 7; Length 777;
Best Local Similarity 84.0%; Pred. No. 2e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAACACCACTGAGCTGGG 25
|||||
DB 544 AAAAAACCACTGCGGTG 520

RESULT 27
US-10-027-632-129760/c
; Sequence 129760, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 129760
; LENGTH: 882
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-129760

Query Match 74.4%; Score 18.6; DB 5; Length 882;
Best Local Similarity 84.0%; Pred. No. 2e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAACACCACTGAGCTGGG 25
|||||
DB 594 AAAAAACCAAACTGACTGGG 570

RESULT 28
US-10-027-632-129761/c
; Sequence 129761, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 129761
; LENGTH: 882
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-129761

Query Match 74.4%; Score 18.6; DB 5; Length 882;
Best Local Similarity 84.0%; Pred. No. 2e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAACACCACTGAGCTGGG 25
|||||
DB 594 AAAAAACCAAACTGACTGGG 570

RESULT 29
US-10-027-632-129762/c
; Sequence 129762, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720

```
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 129762
; LENGTH: 882
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-129762

Query Match          74.4%; Score 18.6; DB 5; Length 882;
Best Local Similarity 84.0%; Pred. No. 2e+02; 4; Indels 0; Gaps 0;
Matches 21; Conservative 0; Mismatches 0;

QY 1 AAAAAACAACCACTGAGCTGGG 25
    |||||
Db 594 AAAAAACAACCACTGAGCTGGG 570

RESULT 30
US-10-027-632-159900/c
; Sequence 159900, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 159900
; LENGTH: 882
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-159900

Query Match          74.4%; Score 18.6; DB 5; Length 882;
Best Local Similarity 84.0%; Pred. No. 2e+02; 4; Indels 0; Gaps 0;
Matches 21; Conservative 0; Mismatches 0;

QY 1 AAAAAACAACCACTGAGCTGGG 25
    |||||
Db 594 AAAAAACAACCACTGAGCTGGG 570

RESULT 31
US-10-027-632-129760/c
; Sequence 129760, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 129760
; LENGTH: 882
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-129760

Query Match          74.4%; Score 18.6; DB 6; Length 882;
Best Local Similarity 84.0%; Pred. No. 2e+02; 4; Indels 0; Gaps 0;
Matches 21; Conservative 0; Mismatches 0;

QY 1 AAAAAACAACCACTGAGCTGGG 25
    |||||
Db 594 AAAAAACAACCACTGAGCTGGG 570

RESULT 32
US-10-027-632-129761/c
; Sequence 129761, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 129761
; LENGTH: 882
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-129761

Query Match          74.4%; Score 18.6; DB 6; Length 882;
Best Local Similarity 84.0%; Pred. No. 2e+02; 4; Indels 0; Gaps 0;
Matches 21; Conservative 0; Mismatches 0;

QY 1 AAAAAACAACCACTGAGCTGGG 25
    |||||
Db 594 AAAAAACAACCACTGAGCTGGG 570

RESULT 33
US-10-027-632-129762/c
; Sequence 129762, Application US/10027632
; Publication No. US20030204075A9
```

```

; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 129762
; LENGTH: 882
; TYPE: DNA
; ORGANISM: Human
; US-10-027-632-129762

Query Match          74.4%; Score 18.6; DB 6; Length 882;
Best Local Similarity 84.0%; Pred. No. 2e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAAACCACTGAGCTGGG 25
Db 594 AAAAAACAACAACTGAGCTGGG 570

RESULT 34
US-10-027-632-159900/c
; Sequence 159900, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 159900
; LENGTH: 882
; TYPE: DNA
; ORGANISM: Human
; US-10-027-632-159900

Query Match          74.4%; Score 18.6; DB 6; Length 882;
Best Local Similarity 84.0%; Pred. No. 2e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

```

Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAAACCACTGAGCTGGG 25
Db 594 AAAAAACAACAACTGAGCTGGG 570

RESULT 35
US-09-925-065A-82610/c
; Sequence 82610, Application US/09925065A
; Publication No. US2005028172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 82610
; LENGTH: 962
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-925-065A-82610

Query Match          74.4%; Score 18.6; DB 4; Length 962;
Best Local Similarity 84.0%; Pred. No. 2e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAAACCACTGAGCTGGG 25
Db 549 AAAAAACAACCACTGAAATGAG 525

RESULT 36
US-10-220-382-32/c
; Sequence 32, Application US/10220382
; Publication No. US2003011911A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: LAL, Preeti
; APPLICANT: TANG, Y. Tom
; APPLICANT: PATTERSON, Chandra
; APPLICANT: YAO, Monique G.
; APPLICANT: SHIH, Leo L.
; APPLICANT: TRIBOULEY, Catherine
; APPLICANT: LU, Dyung Aina M.
; APPLICANT: YUE, Henry
; APPLICANT: KHAN, Farrah A.
; APPLICANT: POLICKY, Jennifer L.
; APPLICANT: AU-YOUNG, Janice
; APPLICANT: YANG, Junming
; APPLICANT: HARLAND, Lee
; APPLICANT: WALSH, Roderick T.
; APPLICANT: LO, Terence P.
; APPLICANT: BOROWSKY, Mark L.
; TITLE OF INVENTION: G-PROTEIN COUPLED RECEPTORS
; FILE REFERENCE: PI-0044 PCT
; CURRENT APPLICATION NUMBER: US/10/220,382
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,854; 60/188,384; 60/190,453; 60/190,730
; PRIOR FILING DATE: 2000-03-03; 2000-03-10; 2000-03-17; 2000-03-20

```

```
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PERL Program
; SEQ ID NO 32
; LENGTH: 1065
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US2003011911A1 7472434CB1
US-10-220-382-32
```

```
Query Match          74.4%; Score 18.6; DB 6; Length 1065;
Best Local Similarity 84.0%; Pred. No. 2.1e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
```

```
QY 1 AAAAAACACCACTGAGCTGGG 25
    ||||| ||||| ||||| |||||
Db 832 AAAAAACCACTGCGGTG 808
```

RESULT 37

```
US-09-965-422-1/c
; Sequence 1, Application US/09965422
; Publication No. US2003021654A1
; GENERAL INFORMATION:
; APPLICANT: Spyttek, Kimberly A
; APPLICANT: Casman, Stacie
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Dickson, Kevin
; APPLICANT: Vernet, Corine
; APPLICANT: Spaderna, Steven K
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Gerlach, Valerie
; APPLICANT: Ellerman, Karen
; APPLICANT: Edinger, Shlomit
; APPLICANT: MacDougall, John R
; APPLICANT: Smithson, Glennda
; APPLICANT: Li, Li
; APPLICANT: Malyankar, Urial M
; APPLICANT: Taylor, Sarah
; APPLICANT: Gunther, Erik
; APPLICANT: Tchernev, Velizar T
; TITLE OF INVENTION: No. US2003021654A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21401-132
; CURRENT APPLICATION NUMBER: US/09/965,422
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: 60/236,286
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: 60/236,284
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: 60/237,581
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/238,735
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/240,736
; PRIOR FILING DATE: 2000-10-16
; PRIOR APPLICATION NUMBER: 60/260,019
; PRIOR FILING DATE: 2001-01-05
; PRIOR APPLICATION NUMBER: 60/260,338
; PRIOR FILING DATE: 2001-01-08
; PRIOR APPLICATION NUMBER: 60/262,156
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/262,498
; PRIOR FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: 60/263,133
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/263,691
; PRIOR FILING DATE: 2001-01-24
; PRIOR APPLICATION NUMBER: 60/266,109
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/271,634
; PRIOR FILING DATE: 2001-02-26
; NUMBER OF SEQ ID NOS: 127
```

```
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1329
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-965-422-1
```

```
Query Match          74.4%; Score 18.6; DB 3; Length 1329;
Best Local Similarity 84.0%; Pred. No. 2.1e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
```

```
QY 1 AAAAAACACCACTGAGCTGGG 25
    ||||| ||||| ||||| |||||
Db 748 AAAAAACCACTGCGGTG 724
```

RESULT 38

```
US-10-012-600B-103
; Sequence 103, Application US/10012600B
; Publication No. US20030064377A1
; GENERAL INFORMATION:
; APPLICANT: Sun, Yongming
; APPLICANT: Recipon, Herve
; APPLICANT: Chen, Sei-Yu
; TITLE OF INVENTION: Compositions and Methods Relating to Prostate Specific Genes and
; FILE REFERENCE: DEX-0265
; CURRENT APPLICATION NUMBER: US/10/012,600B
; PRIOR FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/246,109
; PRIOR FILING DATE: 2000-11-06
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 103
; LENGTH: 1450
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-012-600B-103
```

```
Query Match          74.4%; Score 18.6; DB 5; Length 1450;
Best Local Similarity 84.0%; Pred. No. 2.1e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
```

```
QY 1 AAAAAACACCACTGAGCTGGG 25
    ||||| ||||| ||||| |||||
Db 619 AAAAAACCACTGCGGTG 643
```

RESULT 39

```
US-09-981-566A-43/c
; Sequence 43, Application US/09981566A
; Publication No. US20040005656A1
; GENERAL INFORMATION:
; APPLICANT: Kexuda et al.
; TITLE OF INVENTION: No. US20040005656A1el GPCR-like Proteins and Nucleic Acids Encod
; FILE REFERENCE: 21402-163
; CURRENT APPLICATION NUMBER: US/09/981,566A
; PRIOR FILING DATE: 2001-10-16
; PRIOR APPLICATION NUMBER: 60/240,704
; PRIOR FILING DATE: 2000-10-16
; PRIOR APPLICATION NUMBER: 60/262,159
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/263,340
; PRIOR FILING DATE: 2001-01-22
; PRIOR APPLICATION NUMBER: 60/264,118
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 60/308,203
; PRIOR FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: 60/243,497
; PRIOR FILING DATE: 2000-10-26
; PRIOR APPLICATION NUMBER: 60/244,542
; PRIOR FILING DATE: 2000-10-31
; PRIOR APPLICATION NUMBER: 60/269,031
```

```
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: 60/245,484
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/255,017
; PRIOR FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: 60/263,216
; PRIOR FILING DATE: 2001-01-22
; PRIOR APPLICATION NUMBER: 60/268,225
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 209
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 43
; LENGTH: 2028
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-981-566A-43

Query Match      74.4%; Score 18.6; DB 3; Length 2028;
Best Local Similarity 84.0%; Pred. No. 2.2e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAACACCACTGAGCTGGG 25
Db 1216 AAAAAACCAACCACTGTGCGGTG 1192

RESULT 40
US-09-764-891-6728/c
; Sequence 6728, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6728
; LENGTH: 2464
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-891-6728

Query Match      74.4%; Score 18.6; DB 3; Length 2464;
Best Local Similarity 84.0%; Pred. No. 2.3e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAACACCACTGAGCTGGG 25
Db 714 AAAAAACAAACCACTGAACTGGG 690

RESULT 41
US-09-764-891-6729/c
; Sequence 6729, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6729
; LENGTH: 2464
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-891-6729

Query Match      74.4%; Score 18.6; DB 3; Length 2464;
Best Local Similarity 84.0%; Pred. No. 2.3e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAACACCACTGAGCTGGG 25
Db 714 AAAAAACAAACCACTGAACTGGG 690

RESULT 42
US-10-723-860-7237
; Sequence 7237, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsburg, Wendy M.
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
; FILE REFERENCE: 05882.0193.NFUS01
; CURRENT APPLICATION NUMBER: US/10/723,860
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: 60/429,739
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 8393
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 7237
; LENGTH: 3029
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (705)..(722)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1012)..(1026)
; OTHER INFORMATION: n is a, c, g, or t
US-10-723-860-7237

Query Match      74.4%; Score 18.6; DB 8; Length 3029;
Best Local Similarity 84.0%; Pred. No. 2.3e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAACACCACTGAGCTGGG 25
Db 2946 AAAAAACCAACCACTGAAATGTG 2970

RESULT 43
US-10-172-118-1308
; Sequence 1308, Application US/10172118
; Publication No. US20030224374A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Chris
; APPLICANT: Van 't Veer, Laura
; APPLICANT: Van de Vijver, Marc
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-175-999
; CURRENT APPLICATION NUMBER: US/10/172,118
; CURRENT FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/380,770
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 1308
; LENGTH: 3154
; TYPE: DNA
; ORGANISM: Homo sapiens
```



```
Db 12267 AAAAAAAAAACCAACAGAGCTAGG 12243
||||| ||||| ||||| ||||| ||
RESULT 47
US-10-484-577-681
; Sequence 681, Application US/10484577
; Publication No. US20050032724A1
; GENERAL INFORMATION:
; APPLICANT: EPIDAUROS Biotechnologie Aktiengesellschaft
; TITLE OF INVENTION: Means and methods for improved treatment of cancer based on UGT1A
; FILE REFERENCE: F2285PCT-1
; CURRENT APPLICATION NUMBER: US/10/484,577
; CURRENT FILING DATE: 2004-01-22
; PRIOR APPLICATION NUMBER: PCT/EP 02/08220
; PRIOR FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: EP 01 11 7608.8
; PRIOR FILING DATE: 2001-07-23
; PRIOR APPLICATION NUMBER: EP 02011710.7
; PRIOR FILING DATE: 2002-05-24
; NUMBER OF SEQ ID NOS: 683
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 681
; LENGTH: 128993
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(128993
; LOCATION: )
; OTHER INFORMATION: n=a, c, g or t
US-10-484-577-681
Query Match 74.4%; Score 18.6; DB 8; Length 128993;
Best Local Similarity 84.0%; Pred. No. 3.5e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 AAAAAAAAAACCAACCACTGAGCTGGG 25
||||| ||||| ||||| ||||| ||
Db 121849 AAAAGACACCCCTCACTGAGCTGGG 121873
RESULT 48
US-10-484-577-683/c
; Sequence 683, Application US/10484577
; Publication No. US20050032724A1
; GENERAL INFORMATION:
; APPLICANT: EPIDAUROS Biotechnologie Aktiengesellschaft
; TITLE OF INVENTION: Means and methods for improved treatment of cancer based on UGT1A
; FILE REFERENCE: F2285PCT-1
; CURRENT APPLICATION NUMBER: US/10/484,577
; CURRENT FILING DATE: 2004-01-22
; PRIOR APPLICATION NUMBER: PCT/EP 02/08220
; PRIOR FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: EP 01 11 7608.8
; PRIOR FILING DATE: 2001-07-23
; PRIOR APPLICATION NUMBER: EP 02011710.7
; PRIOR FILING DATE: 2002-05-24
; NUMBER OF SEQ ID NOS: 683
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 683
; LENGTH: 177380
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-484-577-683
Query Match 74.4%; Score 18.6; DB 8; Length 177380;
Best Local Similarity 84.0%; Pred. No. 3.7e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 AAAAAAAAAACCAACCACTGAGCTGGG 25
||||| ||||| ||||| ||||| ||
Db 153700 AAAAGACACCCCTCACTGAGCTGGG 153676
```

```
RESULT 49
US-10-027-632-279949/c
; Sequence 279949, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 279949
; LENGTH: 607
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-279949
Query Match 73.6%; Score 18.4; DB 5; Length 607;
Best Local Similarity 95.0%; Pred. No. 2.4e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 4 AAACAACCAACCACTGAGCTG 23
||||| ||||| ||||| |||||
Db 492 AAACAACCAACCACTGAGCTG 473
RESULT 50
US-10-027-632-279950/c
; Sequence 279950, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 279950
; LENGTH: 607
```

```
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-279950

Query Match      73.6%; Score 18.4; DB 5; Length 607;
Best Local Similarity 95.0%; Pred. No. 2.4e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      4 AAACACACCACCTGAGCTG 23
         |||||
Db      492 AAACACACCACCTGAGCTG 473

Search completed: February 3, 2006, 15:43:28
Job time : 375.556 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 3, 2006, 11:35:31 ; Search time 343.111 Seconds
(without alignments)
61.064 Million cell updates/sec

Title: US-10-719-900-3

Perfect score: 25

Sequence: 1 aaaaacaccaccactgagctggg 25

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 6068529 seqs, 419036697 residues

Total number of hits satisfying chosen parameters: 12137058

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 150 summaries

Database : Published Applications NA New.*

- 1: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
- 2: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
- 5: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
- 7: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
- 9: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
- 10: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
- 11: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	18.6	74.4	1792	7	US-10-750-185-37447
2	18.6	74.4	1792	7	US-10-750-185-37447
3	18.6	74.4	86950	7	US-10-750-185-37447
4	18.4	73.6	1806	7	US-10-750-185-45680
5	18.4	73.6	1806	7	US-10-750-185-45680
6	17.8	71.2	600	8	US-11-136-527-4239
7	17.8	71.2	5383	8	US-11-136-527-4239
8	17.6	70.4	201	7	US-10-995-561-19283
9	17.6	70.4	2817	8	US-11-136-527-1861
10	17.6	70.4	380749	7	US-10-995-561-13216
11	17.4	69.6	21	7	US-10-310-914A-370050
12	17.2	68.8	25	8	US-11-121-849-179004
13	17.2	68.8	1188	7	US-10-750-185-34090
14	17.2	68.8	1188	7	US-10-750-185-34090
15	17.2	68.8	1324	7	US-10-750-185-25618
16	17.2	68.8	1324	7	US-10-750-185-25618
17	17.2	68.8	1826	7	US-10-750-185-27616
18	17.2	68.8	1826	7	US-10-750-185-27616
19	17	68.0	201	7	US-10-995-561-76162
20	17	68.0	201	7	US-10-995-561-76164
21	17	68.0	201	7	US-10-995-561-76165
22	17	68.0	201	7	US-10-995-561-76166

```

96 16.2 64.8 193084 8 US-11-121-086-82
97 16.2 64.8 207835 8 US-11-121-086-39
98 16.2 64.8 207835 8 US-11-121-086-40
99 16.2 64.8 220895 7 US-10-775-169-88
C 100 16 64.0 64 7 US-10-310-914A-20171
C 101 16 64.0 109 7 US-10-310-914A-10751
C 102 16 64.0 201 7 US-10-995-561-27356
C 103 16 64.0 201 7 US-10-995-561-27358
C 104 16 64.0 201 7 US-10-995-561-41476
C 105 16 64.0 201 7 US-10-995-561-41478
C 106 16 64.0 201 7 US-10-995-561-79494
C 107 16 64.0 201 7 US-10-995-561-79495
C 108 16 64.0 201 7 US-10-995-561-79497
C 109 16 64.0 201 7 US-10-995-561-79498
C 110 16 64.0 201 7 US-10-995-561-79499
C 111 16 64.0 201 7 US-10-995-561-79500
C 112 16 64.0 201 7 US-10-995-561-79501
C 113 16 64.0 201 7 US-10-995-561-79502
C 114 16 64.0 201 7 US-10-995-561-79514
C 115 16 64.0 201 7 US-10-995-561-79515
C 116 16 64.0 588 8 US-11-052-554A-666
C 117 16 64.0 600 7 US-10-750-185-24
C 118 16 64.0 600 7 US-10-750-185-1493
C 119 16 64.0 600 7 US-10-750-623-24
C 120 16 64.0 600 7 US-10-750-623-1493
C 121 16 64.0 619 8 US-11-000-463-650
C 122 16 64.0 664 8 US-11-000-463-178
C 123 16 64.0 873 7 US-10-750-185-48965
C 124 16 64.0 873 7 US-10-750-623-48965
C 125 16 64.0 892 7 US-10-750-185-56516
C 126 16 64.0 892 7 US-10-750-623-58516
C 127 16 64.0 1041 7 US-10-750-185-39741
C 128 16 64.0 1041 7 US-10-750-623-39741
C 129 16 64.0 1042 7 US-10-750-185-62101
C 130 16 64.0 1042 7 US-10-750-623-62101
C 131 16 64.0 1148 8 US-11-099-691-24
C 132 16 64.0 1478 7 US-10-750-185-46236
C 133 16 64.0 1478 7 US-10-750-623-46236
C 134 16 64.0 1569 7 US-10-750-185-26134
C 135 16 64.0 1569 7 US-10-750-623-26134
C 136 16 64.0 1762 7 US-10-750-185-58016
C 137 16 64.0 1762 7 US-10-750-623-58016
C 138 16 64.0 1807 7 US-10-750-185-30577
C 139 16 64.0 1807 7 US-10-750-623-30577
C 140 16 64.0 2080 7 US-10-750-185-36859
C 141 16 64.0 2080 7 US-10-750-623-36859
C 142 16 64.0 2346 8 US-11-102-026A-8
C 143 16 64.0 2485 7 US-10-750-185-55627
C 144 16 64.0 2485 7 US-10-750-623-55627
C 145 16 64.0 2489 7 US-10-750-185-61906
C 146 16 64.0 2489 7 US-10-750-623-61906
C 147 16 64.0 2624 7 US-10-750-185-32378
C 148 16 64.0 2624 7 US-10-750-623-32378
C 149 16 64.0 2742 8 US-11-169-041-30
C 150 16 64.0 2743 7 US-10-821-234-267
Sequence 82, Appl
Sequence 39, Appl
Sequence 40, Appl
Sequence 88, Appl
Sequence 20171, A
Sequence 10751, A
Sequence 27356, A
Sequence 27358, A
Sequence 41476, A
Sequence 41478, A
Sequence 79494, A
Sequence 79495, A
Sequence 79497, A
Sequence 79498, A
Sequence 79499, A
Sequence 79500, A
Sequence 79501, A
Sequence 79502, A
Sequence 79514, A
Sequence 79515, A
Sequence 666, App
Sequence 24, Appl
Sequence 1493, Ap
Sequence 1493, Ap
Sequence 650, App
Sequence 178, App
Sequence 48965, A
Sequence 48965, A
Sequence 56516, A
Sequence 56516, A
Sequence 39741, A
Sequence 39741, A
Sequence 62101, A
Sequence 62101, A
Sequence 24, Appl
Sequence 46236, A
Sequence 46236, A
Sequence 26134, A
Sequence 26134, A
Sequence 58016, A
Sequence 58016, A
Sequence 30577, A
Sequence 30577, A
Sequence 36859, A
Sequence 36859, A
Sequence 8, Appl
Sequence 55627, A
Sequence 55627, A
Sequence 61906, A
Sequence 61906, A
Sequence 32378, A
Sequence 32378, A
Sequence 30, Appl
Sequence 267, App
```

ALIGNMENTS

```

RESULT 1
US-10-750-185-37447
; Sequence 37447, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS

RESULT 3
US-10-750-185-37447
; Sequence 37447, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS

RESULT 3
US-10-857-780-5
; Sequence 5, Application US/10857780
; Publication No. US20050272043A1
; GENERAL INFORMATION:
; APPLICANT: ROTH, RICHARD B.
; APPLICANT: BRAUN, ANDREAS
; APPLICANT: KAMMERER, STEFAN M.
; APPLICANT: NELSON, MATTHEW ROBERTS
; APPLICANT: RENEAND, RIKARD HENRY
; APPLICANT: HOYAL-WRIGHTSON, CAROLYN R.
; TITLE OF INVENTION: METHODS FOR IDENTIFYING RISK OF BREAST CANCER AND TREATMENTS
; FILE REFERENCE: SEQ-4069-CP
; CURRENT APPLICATION NUMBER: US/10/857,780
; CURRENT FILING DATE: 2004-05-28

; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 37447
; LENGTH: 1792
; TYPE: DNA
; ORGANISM: Bovine 19866880611864
US-10-750-185-37447

Query Match 74.4%; Score 18.6; DB 7; Length 1792;
Best Local Similarity 84.0%; Pred. No. 37;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAAACCACTGAGCTGGG 25
Db 466 AAAAAAACCACTGAGCTGGG 490

RESULT 2
US-10-750-623-37447
; Sequence 37447, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 37447
; LENGTH: 1792
; TYPE: DNA
; ORGANISM: Bovine 19866880611864
US-10-750-623-37447

Query Match 74.4%; Score 18.6; DB 7; Length 1792;
Best Local Similarity 84.0%; Pred. No. 37;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAAACCACTGAGCTGGG 25
Db 466 AAAAAAACCACTGAGCTGGG 490

RESULT 3
US-10-857-780-5
; Sequence 5, Application US/10857780
; Publication No. US20050272043A1
; GENERAL INFORMATION:
; APPLICANT: ROTH, RICHARD B.
; APPLICANT: BRAUN, ANDREAS
; APPLICANT: KAMMERER, STEFAN M.
; APPLICANT: NELSON, MATTHEW ROBERTS
; APPLICANT: RENEAND, RIKARD HENRY
; APPLICANT: HOYAL-WRIGHTSON, CAROLYN R.
; TITLE OF INVENTION: METHODS FOR IDENTIFYING RISK OF BREAST CANCER AND TREATMENTS
; FILE REFERENCE: SEQ-4069-CP
; CURRENT APPLICATION NUMBER: US/10/857,780
; CURRENT FILING DATE: 2004-05-28
```

Mon Feb 6 12:23:16 2006

PRIOR APPLICATION NUMBER: 10/723,681
PRIOR FILING DATE: 2003-11-25
PRIOR APPLICATION NUMBER: 60/490,234
PRIOR FILING DATE: 2003-07-24
PRIOR APPLICATION NUMBER: 60/525,239
PRIOR FILING DATE: 2003-11-25
NUMBER OF SEQ ID NOS: 4962
SOFTWARE: PatentIn version 3.2
SEQ ID NO 5
LENGTH: 86950
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (20282)..(20282)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (39985)..(39985)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (39989)..(39989)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (51996)..(51996)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (64732)..(64732)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (83916)..(83919)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (86196)..(86196)
OTHER INFORMATION: n is a, c, g, or t
US-10-857-780-5

Query Match 74.4%; Score 18.6; DB 7; Length 86950;
Best Local Similarity 84.0%; Pred. No. 73;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAACAACCACTGAGCTGG 25
||||| ||| ||||| |||||
Db 42395 AAAAAAGATACCACTGGCTGG 42419

RESULT 4
US-10-185-45680/c
; Sequence 45680, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 45680
; LENGTH: 1806

; TYPE: DNA
; ORGANISM: Bovine 19866880824813
US-10-750-185-45680

Query Match 73.6%; Score 18.4; DB 7; Length 1806;
Best Local Similarity 95.0%; Pred. No. 45;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAAAACAACCACTGAG 20
||||| ||| ||||| |||||
Db 192 AAAAAACAACCACTGAG 173

RESULT 5
US-10-750-623-45680/c
; Sequence 45680, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 45680
; LENGTH: 1806
; TYPE: DNA
; ORGANISM: Bovine 19866880824813
US-10-750-623-45680

Query Match 73.6%; Score 18.4; DB 7; Length 1806;
Best Local Similarity 95.0%; Pred. No. 45;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAAAACAACCACTGAG 20
||||| ||| ||||| |||||
Db 192 AAAAAACAACCACTGAG 173

RESULT 6
US-11-136-527-4239
; Sequence 4239, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4239
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-4239

Query Match 71.2%; Score 17.8; DB 8; Length 600;
Best Local Similarity 90.5%; Pred. No. 69;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

: GENERAL INFORMATION:
:
: APPLICANT: Wyeth
:
: APPLICANT: Mounts, William M
:
: TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
:
: FILE REFERENCE: 031896-041000 (AM101086)

```

```

% ORGANISM: Human
% APPLICANT: Benitach, Isaac
% TITLE OF INVENTION: Bioinformatically determined
% FILE OF INVENTION: uses thereof
% FILE REFERENCE: 06087.0200.CPUS01
% CURRENT APPLICATION NUMBER: US/10/310,914A
% CURRENT FILING DATE: 2002-12-06
% NUMBER OF SEQ ID NOS: 1388402
% SOFTWARE: PatentIn version 3.3
% SEQ ID NO 370050
% LENGTH: 21
% TYPE: RNA
% ORGANISM: Human
US-10-310-914A-370050

```



```

Query Match          69.6%; Score 17.4; DB 7; Length 21;
Best Local Similarity 94.7%; Pred. No. 57;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAAAACAACCACTGTA 19
   |||||
Db 19 AAAAAACAACCACTGTA 1

RESULT 12
US-11-121-849-179004
; Sequence 179004, Application US/11121849
; Publication No. US20050272080A1
; GENERAL INFORMATION:
; APPLICANT: John Palma
; TITLE OF INVENTION: Methods of Genetic Analysis of Formalin Fixed Paraffin Embedded S
; FILE REFERENCE: Microarrays
; CURRENT APPLICATION NUMBER: US/11/121,849
; CURRENT FILING DATE: 2005-05-03
; PRIOR APPLICATION NUMBER: 60/567,949
; PRIOR FILING DATE: 2004-05-03
; NUMBER OF SEQ ID NOS: 673904
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 179004
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapien
US-11-121-849-179004

Query Match          68.8%; Score 17.2; DB 8; Length 25;
Best Local Similarity 86.4%; Pred. No. 73;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 AAAAAACAACCACTGAGCTGGG 25
   |||||
Db 2 AAAGCACTCCACTGNACTGGG 23

RESULT 13
US-10-750-185-34090
; Sequence 34090, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 34090
; LENGTH: 1188
; TYPE: DNA
; ORGANISM: Bovine 19866880893114
US-10-750-185-34090

Query Match          68.8%; Score 17.2; DB 7; Length 1188;
Best Local Similarity 86.4%; Pred. No. 1.4e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAACAACCACTGAGCT 22
   |||||
Db 263 AAGAACACCTTCAGCTGAGCT 284

RESULT 14
US-10-750-623-34090
; Sequence 34090, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 34090
; LENGTH: 1188
; TYPE: DNA
; ORGANISM: Bovine 19866880893114
US-10-750-623-34090

Query Match          68.8%; Score 17.2; DB 7; Length 1188;
Best Local Similarity 86.4%; Pred. No. 1.4e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAACAACCACTGAGCT 22
   |||||
Db 263 AAGAACACCTTCAGCTGAGCT 284

RESULT 15
US-10-750-185-25618
; Sequence 25618, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 25618
; LENGTH: 1324
; TYPE: DNA
; ORGANISM: Bovine 19866881355792
US-10-750-185-25618

Query Match          68.8%; Score 17.2; DB 7; Length 1324;
Best Local Similarity 86.4%; Pred. No. 1.5e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAACAACCACTGAGCT 22
   |||||
Db 1224 AAAAAACAACCTTCAGCTGAGCT 1245

RESULT 16

```

```
US-10-750-623-25618
; Sequence 25618, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; PRIOR FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 25618
; LENGTH: 1324
; TYPE: DNA
; ORGANISM: Bovine 19866881355792
US-10-750-623-25618

Query Match      68.8%; Score 17.2; DB 7; Length 1324;
Best Local Similarity 86.4%; Pred. No. 1.5e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAAACACCACTGAGCT 22
   ||||| ||||| ||||| |||||
Db 1224 AAAAAAACCACTGAGCT 1245

RESULT 17
US-10-750-185-27616/c
; Sequence 27616, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 27616
; LENGTH: 1826
; TYPE: DNA
; ORGANISM: Bovine 19866880429771
US-10-750-185-27616

Query Match      68.8%; Score 17.2; DB 7; Length 1826;
Best Local Similarity 86.4%; Pred. No. 1.6e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAAAACCACTGAGCTG 23
   ||||| ||||| ||||| |||||
Db 60 AAAAAACCACTGGGCTG 39

RESULT 18
US-10-750-623-27616/c
; Sequence 27616, Application US/10750623
; Publication No. US20050287531A1
```

```
US-10-750-623-27616
; Sequence 27616, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 76162
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-76162

Query Match      68.0%; Score 17; DB 7; Length 201;
Best Local Similarity 80.0%; Pred. No. 1.3e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAAAACCACTGAGCTGG 25
   ||||| ||||| ||||| |||||
Db 93 AAAAAAACCACTGAGCTGG 69

RESULT 20
US-10-995-561-76164/c
; Sequence 76164, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
```

```
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 76164
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-995-561-76164

Query Match      68.0%; Score 17; DB 7; Length 201;
Best Local Similarity 80.0%; Pred. No. 1.3e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAAAACACCACCTAGCTGGG 25
||||| ||| ||| ||| ||| |||
Db 94 AAAAAAAACAACTTAGCTGGG 70

RESULT 21
US-10-995-561-76165/c
; Sequence 76165, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 76165
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-995-561-76165

Query Match      68.0%; Score 17; DB 7; Length 201;
Best Local Similarity 80.0%; Pred. No. 1.3e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAAAACACCACCTAGCTGGG 25
||||| ||| ||| ||| ||| |||
Db 94 AAAAAAAACAACTTAGCTGGG 70

RESULT 22
US-10-995-561-76166/c
; Sequence 76166, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 76166
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-995-561-76166

Query Match      68.0%; Score 17; DB 7; Length 201;
Best Local Similarity 80.0%; Pred. No. 1.3e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAAAACACCACCTAGCTGGG 25
||||| ||| ||| ||| ||| |||
Db 95 AAAAAAAACAACTTAGCTGGG 71

RESULT 23
US-10-995-561-76170/c
; Sequence 76170, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 76170
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-995-561-76170

Query Match      68.0%; Score 17; DB 7; Length 201;
Best Local Similarity 80.0%; Pred. No. 1.3e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAAAACACCACCTAGCTGGG 25
||||| ||| ||| ||| ||| |||
Db 96 AAAAAAAACAACTTAGCTGGG 72

RESULT 24
US-11-124-368A-3450/c
; Sequence 3450, Application US/11124368A
; Publication No. US20050287559A1
; GENERAL INFORMATION:
; APPLICANT: James J. Devlin
; APPLICANT: May Luke
; TITLE OF INVENTION: Genetic Polymorphisms Associated with
; TITLE OF INVENTION: Vascular Diseases, Methods of Detection and Uses Thereof
; FILE REFERENCE: CL001524
; CURRENT APPLICATION NUMBER: US/11/124,368A
; CURRENT FILING DATE: 2005-05-09
; PRIOR APPLICATION NUMBER: US 60/568,845
; PRIOR FILING DATE: 2004-05-07
; PRIOR APPLICATION NUMBER: US 60/625,936
; PRIOR FILING DATE: 2004-11-09
; NUMBER OF SEQ ID NOS: 21112
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3450
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-11-124-368A-3450

Query Match      68.0%; Score 17; DB 8; Length 201;
Best Local Similarity 80.0%; Pred. No. 1.3e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAAAACACCACCTAGCTGGG 25
||||| ||| ||| ||| ||| |||
Db 95 AAAAAAAATCCACCACTGAGCAGAG 71

RESULT 25
US-10-750-185-349/c
; Sequence 349, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
```



```

; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-11-29
; PRIOR FILING DATE: 2002-11-08
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 38183
; LENGTH: 1424
; TYPE: DNA
; ORGANISM: Bovine 19866880468223
US-10-750-185-38183

Query Match      68.0%; Score 17; DB 7; Length 1424;
Best Local Similarity 80.0%; Pred. No. 1.8e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAAACACCACCTGAGCTGGG 25
   ||| ||||| ||||| ||| |||
Db 1081 AAACAACACCACCTGAAGTGAG 1105

RESULT 30
US-10-750-623-38183
; Sequence 38183, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 38183
; LENGTH: 1424
; TYPE: DNA
; ORGANISM: Bovine 19866880468223
US-10-750-623-38183

Query Match      68.0%; Score 17; DB 7; Length 1424;
Best Local Similarity 80.0%; Pred. No. 1.8e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAAACACCACCTGAGCTGGG 25
   ||| ||||| ||||| ||| |||
Db 1081 AAACAACACCACCTGAAGTGAG 1105

RESULT 31
US-11-000-463-48
; Sequence 48, Application US/11000463
; Publication No. US20050266423A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Chen, Rui-hong
; APPLICANT: Qian, Xiaohong B.
; APPLICANT: Wang, Zhiwei
; APPLICANT: Wehrman, Tom
; APPLICANT: Zhang, Jie
; APPLICANT: Zhou, Ping
; APPLICANT: Cao, Yi-Cheng
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 785CIP4CN

CURRENT APPLICATION NUMBER: US/11/000,463
CURRENT FILING DATE: 2004-11-29
PRIOR APPLICATION NUMBER: 10/291,265
PRIOR FILING DATE: 2002-11-08
PRIOR APPLICATION NUMBER: PCT/US01/02623
PRIOR FILING DATE: 2001-01-25
PRIOR APPLICATION NUMBER: 09/922,279
PRIOR FILING DATE: 2001-08-03
PRIOR APPLICATION NUMBER: 09/491,404
PRIOR FILING DATE: 2000-01-25
PRIOR APPLICATION NUMBER: 09/617,746
PRIOR FILING DATE: 2000-07-17
PRIOR APPLICATION NUMBER: 09/631,451
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: 09/633,870
PRIOR FILING DATE: 2000-09-15
NUMBER OF SEQ ID NOS: 944
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 48
LENGTH: 1612
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (43)..(1464)
US-11-000-463-48

Query Match      68.0%; Score 17; DB 8; Length 1612;
Best Local Similarity 80.0%; Pred. No. 1.9e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAAACACCACCTGAGCTGGG 25
   ||| ||||| ||||| ||| |||
Db 1570 ATAAAGCAACCCACCACTGCCTGGG 1594

RESULT 32
US-11-128-061-1028
; Sequence 1028, Application US/11128061
; Publication No. US20060003958A1
; GENERAL INFORMATION:
; APPLICANT: Melville, Mark W.
; APPLICANT: Charlebois, Timothy S.
; APPLICANT: Mounts, William M.
; APPLICANT: Hann, Louane E.
; APPLICANT: Sinacore, Martin S.
; APPLICANT: Leonard, Mark W.
; APPLICANT: Brown, Eugene L.
; APPLICANT: Miller, Christopher P.
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES RELATED TO OLIGONUCLEOTIDE ARRAYS
; FILE REFERENCE: 01997.027701
; CURRENT APPLICATION NUMBER: US/11/128,061
; CURRENT FILING DATE: 2005-05-11
; PRIOR APPLICATION NUMBER: US 60/570,425
; PRIOR FILING DATE: 2004-05-11
; NUMBER OF SEQ ID NOS: 7285
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1028
; LENGTH: 1623
; TYPE: DNA
; ORGANISM: Cricetus griseus
US-11-128-061-1028

Query Match      68.0%; Score 17; DB 8; Length 1623;
Best Local Similarity 80.0%; Pred. No. 1.9e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAAACACCACCTGAGCTGGG 25
   ||| ||||| ||||| ||| |||
Db 267 AAAGCAAGCAACCCACCAAGAGG 291

```

```
RESULT 33
US-11-128-049-1028
; Sequence 1028, Application US/111128049
; Publication No. US20060010513A1
; GENERAL INFORMATION:
; APPLICANT: Melville, Mark W.
; APPLICANT: Charlebois, Timothy S.
; APPLICANT: Mounts, William M.
; APPLICANT: Hann, Louane E.
; APPLICANT: Sinacore, Martin S.
; APPLICANT: Leonard, Mark W.
; APPLICANT: Brown, Eugene L.
; APPLICANT: Miller, Christopher P.
; TITLE OF INVENTION: OLIGONUCLEOTIDE ARRAYS TO MONITOR GENE EXPRESSION AND METHODS FOR
; FILE REFERENCE: 01997.027700
; CURRENT FILING DATE: 2005-05-11
; PRIOR FILING DATE: 2004-05-11
; NUMBER OF SEQ ID NOS: 7285
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1028
; LENGTH: 1623
; TYPE: DNA
; ORGANISM: Cricetus griseus
US-11-128-049-1028

Query Match      68.0%; Score 17; DB 8; Length 1623;
Best Local Similarity 80.0%; Pred. No. 1.9e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAAACACCACTGAGCTGGG 25
Db 267 AAAGACGACCACTGAGCTGGG 291

RESULT 34
US-11-000-463-159
; Sequence 159, Application US/11000463
; Publication No. US20050266423A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Chen, Rui-hong
; APPLICANT: Qian, Xiaohong B.
; APPLICANT: Wang, Zhiwei
; APPLICANT: Wehrman, Tom
; APPLICANT: Zhang, Jie
; APPLICANT: Zhou, Ping
; APPLICANT: Cao, Yi-Cheng
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 785CIPACN
; CURRENT APPLICATION NUMBER: US/11/000,463
; CURRENT FILING DATE: 2004-11-29
; PRIOR FILING DATE: 2002-11-08
; PRIOR APPLICATION NUMBER: 10/291,265
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: PCT/US01/02623
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: 09/922,279
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 09/491,404
; PRIOR FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: 09/617,746
; PRIOR FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 09/631,451
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 09/633,870
; PRIOR FILING DATE: 2000-09-15
; NUMBER OF SEQ ID NOS: 944
; SOFTWARE: FastSeq for Windows Version 3.0
```

```
; SEQ ID NO 159
; LENGTH: 1644
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (66)..(1499)
US-11-000-463-159

Query Match      68.0%; Score 17; DB 8; Length 1644;
Best Local Similarity 80.0%; Pred. No. 1.9e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAAACACCACTGAGCTGGG 25
Db 1601 ATAAAGCACCACCACTGCCCTGGG 1625

RESULT 35
US-10-775-169-237
; Sequence 237, Application US/10775169
; Publication No. US20050287532A9
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Burczynski, Michael
; APPLICANT: Twine, Natalie
; APPLICANT: Dorner, Andrew
; APPLICANT: Trepicchio, William
; TITLE OF INVENTION: Method for Monitoring Drug Activities In Vivo
; FILE REFERENCE: AM101080 (031896-013000)
; CURRENT APPLICATION NUMBER: US/10/775,169
; CURRENT FILING DATE: 2004-02-11
; NUMBER OF SEQ ID NOS: 5278
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 237
; LENGTH: 1673
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-775-169-237

Query Match      68.0%; Score 17; DB 7; Length 1673;
Best Local Similarity 80.0%; Pred. No. 1.9e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAAACACCACTGAGCTGGG 25
Db 1601 ATAAAGCACCACCACTGCCCTGGG 1625

RESULT 36
US-10-493-909-57
; Sequence 57, Application US/10493909
; Publication No. US20060015969A1
; GENERAL INFORMATION:
; APPLICANT: LARRICK, JAMES W.
; APPLICANT: WYCOFF, KEITH L.
; TITLE OF INVENTION: NOVEL IMMUNOADHESINS FOR TREATING AND PREVENTING TOXICITY
; FILE REFERENCE: 41514-20004.01
; CURRENT APPLICATION NUMBER: US/10/493,909
; CURRENT FILING DATE: 2004-04-26
; PRIOR APPLICATION NUMBER: PCT/US01/13932
; PRIOR FILING DATE: 2001-04-28
; PRIOR APPLICATION NUMBER: 60/200,298
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 101
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 57
; LENGTH: 2028
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-493-909-57
```

050

; PRIOR FILING DATE: 2003-11-11
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 28
; LENGTH: 2551
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: mab TGN1412 heavy chain
US-10-988-207-28

Query Match 68.0%; Score 17; DB 6; Length 2551;
Best Local Similarity 80.0%; Pred. No. 2e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAAACACCACTGAGCTGGG 25
DB 2445 ATAAAGCACCACTGCCCTGGG 2469

RESULT 41
US-11-000-463-630
; Sequence 630, Application US/11000463
; Publication No. US20050266423A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Chen, Rui-hong
; APPLICANT: Qian, Xiaohong B.
; APPLICANT: Wang, Zhiwei
; APPLICANT: Wehrman, Tom
; APPLICANT: Zhang, Jie
; APPLICANT: Zhou, Ping
; APPLICANT: Cao, Yi-Cheng
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 785CIP4CN
; CURRENT APPLICATION NUMBER: US/11/000,463
; PRIOR FILING DATE: 2004-11-29
; PRIOR APPLICATION NUMBER: 10/291,265
; PRIOR FILING DATE: 2002-11-08
; PRIOR APPLICATION NUMBER: PCT/US01/02623
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 09/922,279
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: 09/491,404
; PRIOR FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: 09/617,746
; PRIOR FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 09/631,451
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 09/633,870
; PRIOR FILING DATE: 2000-09-15
; NUMBER OF SEQ ID NOS: 944
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 630
; LENGTH: 3229
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-000-463-630

Query Match 68.0%; Score 17; DB 8; Length 3229;
Best Local Similarity 80.0%; Pred. No. 2.1e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAAACACCACTGAGCTGGG 25
DB 1595 ATAAAGCACCACTGCCCTGGG 1619

RESULT 42
US-11-000-463-631

; Sequence 631, Application US/11000463
; Publication No. US20050266423A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Chen, Rui-hong
; APPLICANT: Qian, Xiaohong B.
; APPLICANT: Wang, Zhiwei
; APPLICANT: Wehrman, Tom
; APPLICANT: Zhang, Jie
; APPLICANT: Zhou, Ping
; APPLICANT: Cao, Yi-Cheng
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 785CIP4CN
; CURRENT APPLICATION NUMBER: US/11/000,463
; PRIOR FILING DATE: 2004-11-29
; PRIOR APPLICATION NUMBER: 10/291,265
; PRIOR FILING DATE: 2002-11-08
; PRIOR APPLICATION NUMBER: PCT/US01/02623
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 09/922,279
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: 09/491,404
; PRIOR FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: 09/617,746
; PRIOR FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 09/631,451
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 09/633,870
; PRIOR FILING DATE: 2000-09-15
; NUMBER OF SEQ ID NOS: 944
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 631
; LENGTH: 3229
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-000-463-631

Query Match 68.0%; Score 17; DB 8; Length 3229;
Best Local Similarity 80.0%; Pred. No. 2.1e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAAACACCACTGAGCTGGG 25
DB 1595 ATAAAGCACCACTGCCCTGGG 1619

RESULT 43
US-10-750-185-41816
; Sequence 41816, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: Denise, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 41816
; LENGTH: 3451
; TYPE: DNA
; ORGANISM: Bovine 19866880792513

US-10-750-185-41816

Query Match 68.0%; Score 17; DB 7; Length 3451;
Best Local Similarity 80.0%; Pred. No. 2.1e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAAACACCACTGAGCTGGG 25

Db 2531 AAAAAACACCTACTCAGAGCTGGG 2555

RESULT 44

US-10-750-623-41816
; Sequence 41816, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 41816
; LENGTH: 3451
; TYPE: DNA
; ORGANISM: Bovine 19866880792513
US-10-750-623-41816

Query Match 68.0%; Score 17; DB 7; Length 3451;
Best Local Similarity 80.0%; Pred. No. 2.1e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAAACACCACTGAGCTGGG 25

Db 2531 AAAAAACACCTACTCAGAGCTGGG 2555

RESULT 45

US-10-775-169-234/c
; Sequence 234, Application US/10775169
; Publication No. US20050287532A9
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Burczynski, Michael
; APPLICANT: Twine, Natalie
; APPLICANT: Dornier, Andrew
; APPLICANT: Trepicchio, William
; TITLE OF INVENTION: Method for Monitoring Drug Activities In Vivo
; FILE REFERENCE: AM101080 (031896-013000)
; CURRENT APPLICATION NUMBER: US/10/775,169
; CURRENT FILING DATE: 2004-02-11
; NUMBER OF SEQ ID NOS: 5278
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 234
; LENGTH: 35962
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-775-169-234

Query Match 68.0%; Score 17; DB 7; Length 35962;
Best Local Similarity 80.0%; Pred. No. 3.2e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAAACACCACTGAGCTGGG 25

Db 9691 AAAAAAACCAAAATTAGCTGGG 9667

RESULT 46

US-10-995-561-13472/c
; Sequence 13472, Application US/10995561
; Publication No. US2005027054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13472
; LENGTH: 50529
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(50529)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1
US-10-995-561-13472

Query Match 68.0%; Score 17; DB 7; Length 50529;
Best Local Similarity 80.0%; Pred. No. 3.4e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAAACACCACTGAGCTGGG 25

Db 3800 AAAAAAACCACTTAGCTGGG 3776

RESULT 47

US-10-775-169-193
; Sequence 193, Application US/10775169
; Publication No. US20050287532A9
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Burczynski, Michael
; APPLICANT: Twine, Natalie
; APPLICANT: Dornier, Andrew
; APPLICANT: Trepicchio, William
; TITLE OF INVENTION: Method for Monitoring Drug Activities In Vivo
; FILE REFERENCE: AM101080 (031896-013000)
; CURRENT APPLICATION NUMBER: US/10/775,169
; CURRENT FILING DATE: 2004-02-11
; NUMBER OF SEQ ID NOS: 5278
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 193
; LENGTH: 110608
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-775-169-193

Query Match 68.0%; Score 17; DB 7; Length 110608;
Best Local Similarity 80.0%; Pred. No. 3.9e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAAACACCACTGAGCTGGG 25

Db 95809 AAAAAACACAAATTAGCCGGG 95833

RESULT 48

US-11-112-908-24
; Sequence 24, Application US/11112908
; Publication No. US20050260659A1
; GENERAL INFORMATION:

```
; APPLICANT: Harris, Cole
; APPLICANT: Davis, Lisa M.
; TITLE OF INVENTION: Breast Cancer Biomarkers
; FILE REFERENCE: 04-164-US
; CURRENT APPLICATION NUMBER: US/11/112,908
; CURRENT FILING DATE: 2005-04-22
; PRIOR APPLICATION NUMBER: US 60/564,758
; PRIOR FILING DATE: 2004-04-23
; PRIOR APPLICATION NUMBER: US 60/575,978
; PRIOR FILING DATE: 2004-06-01
; PRIOR APPLICATION NUMBER: US 60/631,702
; PRIOR FILING DATE: 2004-11-30
; PRIOR APPLICATION NUMBER: US 60/633,826
; PRIOR FILING DATE: 2004-12-07
; NUMBER OF SEQ ID NOS: 511
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 24
; LENGTH: 150314
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-112-908-24

Query Match      68.0%; Score 17; DB 8; Length 150314;
Best Local Similarity 80.0%; Pred. No. 4e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAAACACCACTGAGCTGGG 25
   ||||| ||| ||| ||| |||
DB 97203 AAAAAACAAAAACAATTAGCTGGG 97227

RESULT 49
US-11-172-274-1/c
; Sequence 1, Application US/11/172274
; Publication No. US20060014253A1
; GENERAL INFORMATION:
; APPLICANT: University Of New Hampshire
; APPLICANT: Sower, Stacia A
; APPLICANT: Silver, Matt
; TITLE OF INVENTION: Novel Polynucleotides Encoding Lamprey GnRH-III
; FILE REFERENCE: 9815/59339
; CURRENT APPLICATION NUMBER: US/11/172,274
; CURRENT FILING DATE: 2005-06-30
; PRIOR APPLICATION NUMBER: US/10/170,096
; PRIOR FILING DATE: 2002-06-12
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 790
; TYPE: DNA
; ORGANISM: g. australis
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: (199)..()
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (124)..(405)
; OTHER INFORMATION:
US-11-172-274-1

Query Match      67.2%; Score 16.8; DB 8; Length 790;
Best Local Similarity 90.0%; Pred. No. 2e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAACACCACTGAG 20
   ||||| ||||| |||||
DB 59 AAAAAATCAACCAACTGAG 40

RESULT 50
US-10-775-169-345/c
; Sequence 345, Application US/10775169
```

```
; Publication No. US20050287532A9
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Burczynski, Michael
; APPLICANT: Twine, Natalie
; APPLICANT: Dorner, Andrew
; APPLICANT: Trepicchio, William
; TITLE OF INVENTION: Method for Monitoring Drug Activities In Vivo
; FILE REFERENCE: AM101080 (031896-013000)
; CURRENT APPLICATION NUMBER: US/10/775,169
; CURRENT FILING DATE: 2004-02-11
; NUMBER OF SEQ ID NOS: 5278
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 345
; LENGTH: 128978
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-775-169-345

Query Match      67.2%; Score 16.8; DB 7; Length 128978;
Best Local Similarity 90.0%; Pred. No. 4.8e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAACACCACTGAG 20
   ||||| ||||| ||||| |||
DB 78466 AAAAAACACCACTGAG 78447

Search completed: February 3, 2006, 16:19:47
Job time : 346.111 secs
```

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 3, 2006, 10:53:23 ; Search time 863 Seconds
(without alignments)
1646.682 Million cell updates/sec

Title: US-10-719-900-5
Perfect score: 25
Sequence: 1 aaaaaaacatgaccactgggtgt 25

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues
Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 150 summaries

Database :

GenEmbl.*
1: gb_ba.*
2: gb_in.*
3: gb_env.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pr.*
9: gb_ro.*
10: gb_ets.*
11: gb_ey.*
12: gb_un.*
13: gb_vi.*
14: gb_hcg.*
15: gb_pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
C 1	23.4	93.6	187589	9	AC153565	AC153565 Mus muscu
C 2	23.4	93.6	228194	9	AC153893	AC153893 Mus muscu
C 3	20.8	83.2	57248	8	HS15E1	AL021546 Human DNA
C 4	20.8	83.2	129794	8	AL353812	AL499609 Human DNA
C 5	20.8	83.2	136515	8	AL499609	AC019360 Homo sapi
C 6	20.8	83.2	153605	14	AC019360	AC133867 Mus muscu
C 7	20.8	83.2	168354	14	AC133867	AC142146 Mus muscu
C 8	20.8	83.2	170429	9	AC142146	AL590454 Homo sapi
C 9	20.8	83.2	171970	14	AL590454	AC025988 Homo sapi
C 10	20.8	83.2	177654	14	AC025988	Z97199 Homo sapien
C 11	20.8	83.2	183861	14	HS75N14	AC129236 Homo sapi
C 12	20.8	83.2	187546	14	AC129236	BV293157 S232P6532
C 13	20.2	80.8	601	10	BV293157	AC165059 Phakopsor
C 14	20.2	80.8	40745	14	AC165059	AC100243 Mus muscu
C 15	20.2	80.8	57175	14	AC100243	AC111185 Homo sapi
C 16	20.2	80.8	62356	14	AC111185	AL356141 Human DNA
C 17	20.2	80.8	87871	8	AL356141	AC154170 Alligator
C 18	20.2	80.8	134231	14	AC154170	

AF215848	Homo sapi	AF215848	14	199451	14	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14	18.8	75.2	199451	14
----------	-----------	----------	----	--------	----	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----	------	------	--------	----

Sequence similarities were identified using Powerblast by Jinghui Zhang.

Exon/Intron boundaries of identified genes were chosen if there were canonical splice junctions that maintained sequence continuity across the splice junctions.

FEATURES

```

source
1. .187589
   /organism="Mus musculus"
   /mol_type="genomic DNA"
   /db_xref="taxon:10090"
   /chromosome="10"
   /clones="RP23-228H2"

misc_feature
1. .44609
   /notes="overlaps bases 183586. .228194 of clone AC153893"
   /function="Clone overlap"

repeat_region
135. .225
   /rpt_family="PB1D7"

repeat_region
317. .555
   /rpt_family="B4A"

repeat_region
614. .656
   /rpt_family="CA)n"

repeat_region
1502. .1538
   /rpt_family="CAG)n"

repeat_region
1894. .2386
   /rpt_family="L1MB4"

repeat_region
2850. .2884
   /rpt_family="TAAA)n"

repeat_region
3563. .4822
   /rpt_family="ORR1D-int"

repeat_region
4823. .4866
   /rpt_family="(TG)n"

repeat_region
4867. .4892
   /rpt_family="ORR1D-int"

repeat_region
4893. .5214
   /rpt_family="ORR1D"

repeat_region
5768. .5823
   /rpt_family="(GGAA)n"

repeat_region
5824. .5983
   /rpt_family="(GGGAA)n"

repeat_region
6289. .6509
   /rpt_family="B3"

repeat_region
7086. .7127
   /rpt_family="(CA)n"

repeat_region
8477. .8497
   /rpt_family="AT_rich"

repeat_region
complement(8550. .8740)
   /rpt_family="B3A"

repeat_region
9885. .9929
   /rpt_family="(GGGA)n"

repeat_region
complement(11090. .11407)
   /rpt_family="ORR1B1"

repeat_region
complement(11553. .11706)
   /rpt_family="RSINE1"

repeat_region
complement(11717. .11823)
   /rpt_family="PB1D10"

repeat_region
complement(11744. .11843)
   /rpt_family="MT2B"

repeat_region
14853. .14997
   /rpt_family="B4A"

repeat_region
16135. .16413
   /rpt_family="B4"

repeat_region
16443. .16496
   /rpt_family="(CA)n"

repeat_region
17042. .17267
   /rpt_family="MTD"

repeat_region
18399. .18620
   /rpt_family="AT_rich"

repeat_region
21422. .21471
   /rpt_family="(TG)n"

repeat_region
21626. .21658
   /rpt_family="(TCC)n"

repeat_region
complement(21684. .21852)

```

```

repeat_region
22693. .22785
   /rpt_family="B4"

repeat_region
22976. .23025
   /rpt_family="B1F"

repeat_region
23121. .23171
   /rpt_family="B4A"

repeat_region
24199. .24296
   /rpt_family="(CATA)n"

repeat_region
24348. .24369
   /rpt_family="PB1D10"

repeat_region
25089. .25169
   /rpt_family="AT_rich"

repeat_region
25170. .25227
   /rpt_family="B1_Mur2"

repeat_region
25230. .25388
   /rpt_family="(GAAA)n"

repeat_region
26120. .26214
   /rpt_family="(GGAA)n"

repeat_region
26335. .26517
   /rpt_family="B1_Mur1"

repeat_region
26574. .26761
   /rpt_family="B3"

repeat_region
complement(26775. .26989)
   /rpt_family="L1MB8"

repeat_region
27249. .27389
   /rpt_family="B3"

repeat_region
27801. .28024
   /rpt_family="(CATATA)n"

repeat_region
28461. .28517
   /rpt_family="B4A"

repeat_region
complement(29233. .29348)
   /rpt_family="(CA)n"

repeat_region
29701. .29831
   /rpt_family="MIR"

repeat_region
complement(29831. .29831)
   /rpt_family="B1F"

repeat_region
30002. .30040
   /rpt_family="A-rich"

repeat_region
30090. .30240
   /rpt_family="B1_Mur3"

repeat_region
30581. .30632
   /rpt_family="(CA)n"

repeat_region
31143. .31165
   /rpt_family="AT_rich"

repeat_region
31166. .31276
   /rpt_family="AT_rich"

repeat_region
31279. .31438
   /rpt_family="PB1D10"

repeat_region
complement(31438. .31438)
   /rpt_family="B3A"

```

Query Match 93.6%; Score 23.4; DB 9; Length 187589;
 Best Local Similarity 96.0%; Pred. No. 3.4;
 Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AAAAAACATGACCAACTGGGCTGT 25
 |||||
 Db 32277 AAAAAACATGACCAACTGGGCTGT 32253

RESULT 2

AC153893/c
 LOCUS AC153893 228194 bp DNA linear ROD 29-JAN-2005
 DEFINITION Mus musculus 10 BAC RP23-124O20 (Roswell Park Cancer Institute
 (C57BL/6J Female) Mouse BAC Library) complete sequence.
 ACCESSION AC153893
 VERSION AC153893.3 GI:58257766
 KEYWORDS HTG.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muroidea; Muridae; Murinae; Mus.
 1 (bases 1 to 228194)
 Muzny, D., Adams, C., Agbai II, O., Allen, C., Alabrooks, S., Archer, P.,
 Arredondo, H., Bandaranaike, D., Bangura, L., Beltran, B., Beltran, R.,

Berauducci, A., Biswalo, K., Blyth, P., Bonham, H., Buhay, C., Burch, P., Cadoree, I., Canada, A., Cardenas, V., Carter, K., Cavazos, I., Chacko, J., Chahrour, M., Chavez, D., Chen, A., Chen, G., Chen, R., Cheng, M.-T., Chu, J., Clerck, K., Cockrell, R., Coyle, M., Cree, A., Curry, S., Dai, W., Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Donlin, J., McCauley, S., Dugan-Rocha, S., Dunn, A., Durbin, K., Dziuda, D., Egan, A., Escotto, M., Espinosa, V., Eugene, C., Fa, M., Fernandez, S., Fernando, P., Flagg, N., Forbes, L., Foster, P., Fowler, G., Fu, Q., Fuh, E., Garcia, A., Garcia, N., Garner, T., Gaskin, C., Gench, S., Ghose, S., Gill, R., Gonzalez, D., Gonzalez-Garay, M., Guevara, W., Holder, M., Haaland, W., Haerlen, K., Hall, B., Hamid, H., Hamilton, K., Harbes, B., Harris, R., Havlak, P., Hawes, A., Hawkins, E., Hayes, S., Hemphill, L., Hernandez, J., Hines, S., Hitchens, M., Hodgson, A., Hogue, M., Hollings, B., Howell, L. T., Hulyk, S., Hume, J., Imo, K., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Kalafus, K., Kelly, S., Keys, T., Khan, Z., King, L., Kovar, C., Kowis, A., Kowis, C., Lara, F., Leal, S., Lee, K., Lee, S., LeGall, F. I., Lemon, S., Lewis, L., Li, B., Li, Y., Li, Z., Linnell, M., Liu, W., Liu, Y.-S., Liu, Y., Liyanage, D., London, P., Lopez, J., Lorensuhewa, L., Lozado, R., Luk, T., Madu, R., Maheshwari, M., Mahoney, C., Malloy, K., Mansouri, D., Martinez, E., McClelland, H., McPherson, J., Mercadao, C., Metzker, M., Milosavljevic, A., Minja, E., Morgan, M., Morris, S., Munidasa, M., Murray, D., Nazareth, L., Ngo, D., Nguyen, N., Norwig-Eastaugh, E., Nott, A., Nwaokeme, O., Obregon, M., Ochi-Okorie, C., Odeh, E., Okwuonu, G., Okwuonu, K., Parker, D., Pasternak, S., Patel, B., Patel, V., Paul, H., Perez, A., Perez, L., Petrosino, J., Pham, T., Primus, E., Pu, L.-L., Puazo, M., Qin, X., Quinn, A., Quiroz, J., Rabata, D., Rachlin, E., Raigh, R., Ren, Y., Reuter, M., Richards, S., Rives, C., Rodriguez, F., Rojas, A., Ruiz, S. J., Sana, M., Sanders, W., Santibanez, J., Santos, R., Savery, G., Scherer, S., Shen, H., Shen, Y., Sisson, I., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Svatek, A., Taylor, E., Taylor, T., Thomas, N., Thorn, R., Thornton, R., Trejos, Z., Usmami, K., Vargo, C., Verdusco, D., Villaseana, D., Vark, D., Volkov, A., Waldron, L., Walker, B., Wang, Q., Wang, S., Warren, J., Wei, X., Wheeler, D., Williams, G., Williams, R., Worley, K., Wright, R., Wu, J., Yakub, S., Yan, K., Yuan, Y., Yu, F., Zhang, J., Zhang, L., Zhang, Z., Zhou, J., Weinstock, G. and Gibbs, R.

Direct Submission
Unpublished
2 (bases 1 to 228194)
Worley, K.C.

Direct Submission
Submitted (18-DEC-2004) Human Genome Sequencing Center, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 228194)
Worley, K.C.

Direct Submission
Submitted (08-JAN-2005) Human Genome Sequencing Center, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
4 (bases 1 to 228194)
Worley, K.C.

Direct Submission
Submitted (28-JAN-2005) Human Genome Sequencing Center, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
5 (bases 1 to 228194)
Worley, K.C.

Direct Submission
Submitted (29-JAN-2005) Human Genome Sequencing Center, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Jan 28, 2005 this sequence version replaced gi:57334844.
Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality does not meet this standard, it will be indicated in the annotation.

The repeat regions shown were identified using RepeatMasker by Adrian Smit.

Sequence similarities were identified using Powerblast by Jinghui Zhang.

Exon/Intron boundaries of identified genes were chosen if there were canonical splice junctions that maintained sequence continuity across the splice junctions.

FEATURES
source
location/Qualifiers
1..228194
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="RP23-124020"
1..1267
/rpt_family="L1Md_T"
complement(5200..5414)
/rpt_family="URR1A"
5482..5658
/rpt_family="TTC)n"
5573
/note="polymorphism: T/C, MSZWP/outside reads"
/function="polymorphic site"
5659..5762
/rpt_family="TCC)n"
5771..5868
/rpt_family="B3"
6322..6357
/rpt_family="GA)n"
6857..6912
/rpt_family="GAAA)n"
8508..8530
/rpt_family="AT_rich"
10134..10232
/rpt_family="CA)n"
10776..11064
/rpt_family="B4A"
11081..11138
/rpt_family="CA)n"
complement(12651..12874)
/rpt_family="URR1A"
13728..13789
/rpt_family="GGGA)n"
14331..14353
/rpt_family="TGGG)n"
complement(15471..15529)
/rpt_family="RMR30"
16544..16678
/rpt_family="GGGAGA)n"
16738..16791
/rpt_family="CA)n"
17125..17440
/rpt_family="ORR1B1"
17477..17607
/rpt_family="RSINE1"
17613..17654
/rpt_family="CA)n"
18510..18656
/rpt_family="RSINE1"
19893..19998
/rpt_family="PB1D10"
20005..20063
/rpt_family="A-rich"
20782..20976
/rpt_family="URR1A"
21547..21827
/rpt_family="B4A"
complement(23063..23205)
/rpt_family="B1_Mus1"
25349..25375
/rpt_family="TA)n"
complement(26639..26848)
/rpt_family="B3"
28245..28498
/rpt_family="B4A"
complement(29029..29107)
/rpt_family="B4A"

```

repeat_region 29544..29579
/rpt_family="AT_rich"
repeat_region 29582..29657
/rpt_family="ID4_"
/rpt_family="ID4_"
repeat_region 29790..29822
/rpt_family="(GA)n"
repeat_region 29828..29927
/rpt_family="(GGA)n"
repeat_region 30730..30848
/rpt_family="PB1D9"
repeat_region 30889..30980
/rpt_family="B1_Mus1"
repeat_region 30959..31002
/rpt_family="ID_B1"
repeat_region 31018..31117
/rpt_family="(GGGGA)n"
repeat_region complement(31293..31412)
/rpt_family="B1_Mur3"
repeat_region complement(32054..32268)
/rpt_family="L1MB7"
repeat_region complement(33129..33376)
/rpt_family="Lx6"
repeat_region 34113..34143
/rpt_family="AT_rich"
repeat_region complement(34145..34346)
/rpt_family="B3"
repeat_region 35847..35879
/rpt_family="AT_rich"
repeat_region 37543..37641
/rpt_family="MLT1B"
repeat_region 38638..38770
/rpt_family="L1MA9"
repeat_region 38805..38843
/rpt_family="(CA)n"
repeat_region 38844..38891
/rpt_family="(GA)n"
repeat_region 38916..39185
/rpt_family="L1MA9"
repeat_region 39194..39305
/rpt_family="PB1"

```

Query Match 93.6%; Score 23.4; DB 9; Length 228194;
 Best Local Similarity 96.0%; Pred. No. 3.4;
 Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

Qy 1 AAAAAAATGACCACTGGGCTGT 25
|||||
Db 215862 AAAAAAATGACCACTGGGCTGT 215838
|||||

```

```

RESULT 3
HS15E1 57248 bp DNA linear PRI 16-APR-2005
LOCUS Human DNA sequence from clone XX-15E1 on chromosome 12, complete
DEFINITION
ACCESSION AL021546
VERSION AL021546.1 GI:2826890
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 57248)
MURPHY.L.
Direct Submission
TITLE Submitted (05-JUN-2003) Wellcome Trust Sanger Institute, Hinxton,
JOURNAL Cambridgehire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
----- Genome Center
COMMENT Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk

```

Contact: humquery@sanger.ac.uk

 During sequence assembly data is compared from overlapping clones.
 Where differences are found these are annotated as variations
 together with a note of the overlapping clone name. Note that the
 variation annotation may not be found in the sequence submission
 corresponding to the overlapping clone, as we submit sequences with
 only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all
 regions were either double-stranded or sequenced with an alternate
 chemistry or covered by high quality data (i.e., phred quality >=
 30); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by at least
 one plasmid subclone or more than one M13 subclone; and the
 assembly was confirmed by restriction digest, except on the rare
 occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession
 numbers given in the feature table with their source databases:
 En:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
 on the WORMPEP database can be found at

http://www.sanger.ac.uk/Projects/C_elegans/wormpep/IMPORTANT: This
 sequence is not the entire insert of clone XX-15E1 It may be
 shorter because we sequence overlapping sections only once, except
 for a short overlap.

The true left end of clone RP1-75N14 is at 55991 in this sequence.
 The true right end of clone XX-166H1 is at 1098 in this sequence.

FEATURES

Source	Location/Qualifiers
1..57248	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/db_xref="taxon:9606"
	/chromosome="12"
	/clone="XX-15E1"
1..97	/note="L1MC1 repeat: matches 6233..6324 of consensus"
132..438	/note="AluYa8 repeat: matches 1..302 of consensus"
438..718	/note="AluJo repeat: matches 1..281 of consensus"
745..1035	/note="AluX repeat: matches 1..292 of consensus"
1039..1335	/note="AluX repeat: matches 3..299 of consensus"
1447..1552	/note="MIR repeat: matches 100..212 of consensus"
1830..2117	/note="AluSq repeat: matches 1..291 of consensus"
2118..2420	/note="AluY repeat: matches 8..307 of consensus"
2748..2976	/note="MIR repeat: matches 28..262 of consensus"
3027..3130	/note="52 copies 2 mer cc 60% conserved"
3215..3507	/note="AluY repeat: matches 1..307 of consensus"
3551..3862	/note="AluX repeat: matches 1..312 of consensus"
3921..4018	/note="L2 repeat: matches 2597..2706 of consensus"
4036..4342	/note="AluJo repeat: matches 1..292 of consensus"
complement(4343..4423)	/note="match: GSS: Em:AQ621351"
4424..4487	/note="MER5A repeat: matches 45..116 of consensus"
4573..4874	/note="AluX repeat: matches 1..306 of consensus"
4948..5076	/note="FLAM C repeat: matches 5..133 of consensus"
5082..5150	/note="L2 repeat: matches 2404..2471 of consensus"
5713..5901	/note="MIR repeat: matches 13..194 of consensus"

```
repeat_region 5910. .6217
/note="LJMC5 repeat: matches 7444. .7758 of consensus"
repeat_region 6218. .6245
/note="14 copies 2 mer gt 89% conserved"
repeat_region 6247. .6411
/note="AluJo repeat: matches 109. .285 of consensus"
repeat_region 6412. .6716
/note="AluJ repeat: matches 2. .306 of consensus"
repeat_region 6722. .7039
/note="AluSx repeat: matches 1. .311 of consensus"
repeat_region 7040. .7165
/note="AluJo repeat: matches 11. .105 of consensus"
repeat_region 7183. .7335
/note="LJMA4A repeat: matches 5998. .6277 of consensus"
repeat_region 7371. .7592
/note="AluJb repeat: matches 1. .228 of consensus"
repeat_region 7669. .7759
/note="L2 repeat: matches 2629. .2714 of consensus"
repeat_region 7760. .8058
/note="AluY repeat: matches 2. .297 of consensus"
repeat_region 8059. .8087
/note="Alu repeat: matches 117. .145 of consensus"
repeat_region 8088. .8161
/note="AluS repeat: matches 231. .304 of consensus"
repeat_region 8179. .8481
/note="AluSp repeat: matches 1. .304 of consensus"
repeat_region 8489. .8771
/note="AluJo repeat: matches 1. .288 of consensus"
repeat_region 8990. .9055
/note="33 copies 2 mer ta 72% conserved"
repeat_region 9057. .9219
/note="AluSg/x repeat: matches 134. .296 of consensus"
repeat_region 9221. .9308
/note="L2 repeat: matches 2637. .2725 of consensus"
repeat_region 9339. .9596
/note="L2 repeat: matches 2135. .2419 of consensus"
repeat_region 9601. .9735
/note="AluJb repeat: matches 1. .135 of consensus"
repeat_region 9736. .9926
/note="AluSx repeat: matches 1. .210 of consensus"
repeat_region 9927. .10144
/note="AluSx repeat: matches 1. .216 of consensus"
repeat_region 10145. .10252
/note="AluSx repeat: matches 210. .312 of consensus"
repeat_region 10483. .10556
/note="L2 repeat: matches 2624. .2695 of consensus"
repeat_region 10601. .10894
/note="AluY repeat: matches 1. .297 of consensus"
repeat_region 10895. .11031
/note="AluJb repeat: matches 1. .137 of consensus"
repeat_region 11068. .11200
/note="FLAM C repeat: matches 1. .133 of consensus"
repeat_region 11525. .11682
/note="MIR repeat: matches 92. .262 of consensus"
repeat_region 11733. .12027
/note="AluSg repeat: matches 1. .295 of consensus"
repeat_region 12027. .12306
/note="AluJb repeat: matches 1. .281 of consensus"
repeat_region 12426. .12723
/note="AluSx repeat: matches 1. .298 of consensus"
repeat_region 13132. .13335
/note="MIR repeat: matches 6. .212 of consensus"
repeat_region 13445. .13594
/note="AluSg/x repeat: matches 155. .297 of consensus"
repeat_region 13595. .13895
/note="AluSx repeat: matches 1. .292 of consensus"
repeat_region 13896. .14197
/note="AluSx repeat: matches 1. .305 of consensus"
repeat_region 14198. .14222
/note="AluSg/x repeat: matches 113. .155 of consensus"
repeat_region 14488. .14785
/note="AluSg repeat: matches 3. .300 of consensus"
repeat_region 14804. .14854

/note="MIR repeat: matches 210. .260 of consensus"
14859. .14925
/note="MIR repeat: matches 2. .69 of consensus"
14926. .15105
/note="AluJo repeat: matches 1. .183 of consensus"
15111. .15398
/note="AluSx repeat: matches 6. .292 of consensus"
15466. .15525
/note="MIR repeat: matches 14. .85 of consensus"
15526. .15642
/note="FLAM A repeat: matches 20. .136 of consensus"
15858. .15949
/note="MIR repeat: matches 23. .122 of consensus"
15967. .16270
/note="AluSx repeat: matches 1. .305 of consensus"
16271. .16571
/note="AluJb repeat: matches 1. .300 of consensus"
16720. .16843
/note="FLAM C repeat: matches 7. .134 of consensus"
16877. .16902
/note="13 copies 2 mer aa 100% conserved"
16903. .17218
/note="AluSg repeat: matches 1. .313 of consensus"
17307. .17556
/note="L2 repeat: matches 2485. .2750 of consensus"
17895. .20523
/gene="COX6A2"
Join(17895. .18023,18173. .18315,20248. .20523)
/gene="COX6A2"
/product="15E1.4 (Cytochrome C Oxidase Polypeptide
Vfa-liver precursor (EC 1.9.3.1))"
/evidence=not experimental
Join(17921. .18023,18173. .18315,20248. .20331)

CDS
Query Match 83.2%; Score 20.8; DB 8; Length 57248;
Best Local Similarity 91.7%; Pred. No. 63;
Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAACAATGACCACTGGGCTG 24
Db 42908 AAAAAACAATGTCACACTGGTCTG 42931

RESULT 4
AL353812/c AL353812 129794 bp DNA linear PRI 18-MAY-2005
LOCUS Human DNA sequence from clone RP11-384D7 on chromosome 20 Contains
DEFINITION the 5' end of the C20orf22 for chromosome 20 open reading frame 22
(DKFZp434P106), the PPIAP2 gene for peptidylprolyl isomerase A
(cyclophilin A) pseudogene 2, the 5' end of the gene for a novel
protein (K1AA0186) and three CpG islands, complete sequence.
ACCESSION AL353812
VERSION AL353812.13 GI:8670913
KEYWORDS HTG; C20orf22; CpG island; cyclophilin; DKFZ434P106; K1AA0186;
peptidylprolyl isomerase; PPIAP2.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
REFERENCE 1 (bases 1 to 129794)
AUTHORS Dunn, M.
TITLE Direct Submission
JOURNAL Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vegas@sanger.ac.uk
Clone requests: clonerequest@sanger.ac.uk
On Jun 22, 2000 this sequence version replaced gi:8670605.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
```


Boguslavskiy, L., Boukhgalter, B., Brown, A., Burkett, G.,
 Campopiano, A., Castle, A., Choellapel, Y., Colangelo, M., Collins, S.,
 Callimore, A., Cooke, P., DeAtellano, K., Dewar, K., Diaz, J.S.,
 Dodge, S., Domino, M., Doyle, M., Ferreira, P., FitzHugh, W., Gage, D.,
 Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,
 Grand-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L.,
 Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,
 Klein, J., LaRoque, K., Lamazares, R., Landers, T., Lehocsky, J.,
 Levine, R., Lieu, C., Liu, G., Locke, K., Macdonald, P., Marquis, N.,
 McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPheeters, R.,
 McElmrim, J., Meneus, L., Mhova, T., Miranda, C., Mienga, V., Morrow, J.,
 Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,
 O'Neil, D., Olivier, T.M., Oliver, J., Peterson, K., Pierre, N.,
 Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,
 Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,
 Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
 Tesfaye, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J.,
 Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J.,
 Young, G., Zainoun, J., Zimmer, A. and Zody, M.

Direct Submission
 Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Apr 8, 2000 this sequence version replaced gi:6715815.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center -----
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: <http://www-seq.wi.mit.edu>
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information -----
 Center project name: L5690
 Center clone name: 777_M_5
 ----- Summary Statistics -----
 Sequencing vector: M13; M77815; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.960731
 Consensus quality: 138398 bases at least Q40
 Consensus quality: 145295 bases at least Q30
 Consensus quality: 147783 bases at least Q20
 Insert size: 171000; agarose-fp
 Insert size: 150305; sum-of-contigs
 Quality coverage: 3.8 in Q20 bases; agarose-fp
 Quality coverage: 4.3 in Q20 bases; sum-of-contigs

TITLE
 JOURNAL
 COMMENT

15948 16047: gap of 100 bp
 16048 17790: contig of 1743 bp in length
 17791 17890: gap of 100 bp
 17891 20931: contig of 3041 bp in length
 20932 21031: gap of 100 bp
 21032 24518: contig of 3487 bp in length
 24519 24619: gap of 100 bp
 24619 27964: contig of 3346 bp in length
 27965 28064: gap of 100 bp
 28065 30804: contig of 2740 bp in length
 30805 30904: gap of 100 bp
 30905 34188: contig of 3284 bp in length
 34189 34288: gap of 100 bp
 34289 38917: contig of 4629 bp in length
 38918 43237: gap of 100 bp
 39018 43337: contig of 4220 bp in length
 43338 47566: contig of 4229 bp in length
 47567 47666: gap of 100 bp
 47667 51301: contig of 3635 bp in length
 51302 51401: gap of 100 bp
 51402 55847: contig of 4446 bp in length
 55848 60782: gap of 100 bp
 60783 60882: contig of 4835 bp in length
 60883 65010: contig of 4128 bp in length
 65011 67924: contig of 2814 bp in length
 67925 72012: contig of 3988 bp in length
 72013 72112: gap of 100 bp
 72113 78507: contig of 6395 bp in length
 78508 85236: contig of 6629 bp in length
 85237 91732: contig of 6396 bp in length
 91733 91832: gap of 100 bp
 91833 98668: contig of 6836 bp in length
 98669 107846: contig of 9078 bp in length
 107847 107946: gap of 100 bp
 107947 116112: contig of 8166 bp in length
 116113 124989: contig of 8777 bp in length
 124990 125089: gap of 100 bp
 125090 135741: contig of 10652 bp in length
 135742 135841: gap of 100 bp
 135842 153605: contig of 17764 bp in length.

FEATURES
 source
 1. .153605
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="12"
 /map="12"
 /clone="RP11-777M5"
 /clone_lib="RPC1-11 Human Male BAC"
 1. .1357
 /note="assembly_fragment"
 1358. .1457
 /estimated_length=100
 1458. .2566
 /note="assembly_fragment"
 2567. .2666
 /estimated_length=100
 2667. .3828
 /note="assembly_fragment"
 3829. .3928
 /estimated_length=100
 3929. .5157
 /note="assembly_fragment"
 5158. .5257
 /estimated_length=100

misc_feature
 gap
 misc_feature
 gap
 misc_feature
 gap
 misc_feature
 gap

NOTE: This is a 'working draft' sequence. It currently
 consists of 34 contigs. The true order of the pieces
 is not known and their order in this sequence record is
 arbitrary. Gaps between the contigs are represented as
 runs of N, but the exact sizes of the gaps are unknown.
 This record will be updated with the finished sequence
 as soon as it is available and the accession number will
 be preserved.

1 1357: contig of 1357 bp in length
 1358 1457: gap of 100 bp
 1458 2566: contig of 1109 bp in length
 2567 2666: gap of 100 bp
 2667 3828: contig of 1162 bp in length
 3829 3928: gap of 100 bp
 3929 5157: contig of 1229 bp in length
 5158 5257: gap of 100 bp
 5258 6995: contig of 1738 bp in length
 6996 7095: gap of 100 bp
 7096 8163: contig of 1068 bp in length
 8164 8263: gap of 100 bp
 8264 9738: contig of 1475 bp in length
 9739 9839: gap of 100 bp
 9839 12015: contig of 2177 bp in length
 12016 12115: gap of 100 bp
 12116 14153: contig of 2038 bp in length
 14154 14253: gap of 100 bp
 14254 15947: contig of 1694 bp in length

```

misc_feature      5258..6995
/note="assembly_fragment"
gap              6996..7095
/estimated_length=100
misc_feature      7096..8163
/note="assembly_fragment"
gap              8164..8263
/estimated_length=100
misc_feature      8264..9738
/note="assembly_fragment"
gap              9739..9838
/estimated_length=100
misc_feature      9839..12015
/note="assembly_fragment"
gap              12016..12115

Query Match      83.2%; Score 20.8; DB 14; Length 153605;
Best Local Similarity 91.7%; Pred. No. 61;
Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAACAATGACCACTGGGCTG 24
||||| ||||| ||||| ||||| |||||
Db 103768 AAAAAACAATGTCCAACTGGTCTG 103745

RESULT 7
AC133867 169354 bp DNA linear HTG 04-MAR-2003
LOCUS Mus musculus clone RP24-211L7, WORKING DRAFT SEQUENCE, 11 unordered
DEFINITION
AC133867
AC133867 GI:28827899
VERSION HTG; HTGS PHASE1; HTGS DRAFT.
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Murioidea; Muridae; Murinae; Mus.
1 (bases 1 to 169354)
Birren,B., Nusbaum,C. and Lander,E.
Mus musculus, clone RP24-211L7
Unpublished
2 (bases 1 to 169354)
Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B.,
Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J.,
Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
Liu,G., Maclean,C., Macdonald,P., Major,J., Matthews,C.,
McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Roy,A., Schauer,S., Schupback,R., Seaman,S., Severy,P.,
Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
Tefaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (19-SEP-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 169354)
Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
Boguslavskiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y.,
Collymore,A., Cooke,A., Cooke,P., Corum,B., DeArellano,K.,
Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,
Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S.,
Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,
Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,

```

```

Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
Lindblad-Toh,K., Liu,G., Lui,A., Mabbitt,R., Maclean,C.,
Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J.,
Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
Roman,J., Roy,A., Schauer,S., Schupback,R., Seaman,S., Severy,P.,
Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,
Talamas,J., Tefaye,S., Theodore,J., Topham,K., Travers,M.,
Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (04-MAR-2003) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 4, 2003 this sequence version replaced gi:28416243.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L21215
Center clone name: 211.L.7
----- Summary Statistics
Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 165421 bases at least Q40
Consensus quality: 166530 bases at least Q30
Consensus quality: 166997 bases at least Q20
Insert size: 172000; agarose-1p
Insert size: 167354; sum-of-contigs
Quality coverage: 11.1 in Q20 bases; agarose-1p
Quality coverage: 11.5 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 11 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 20209: contig of 20209 bp in length
* 20210 20309: gap of 100 bp
* 20310 21148: contig of 839 bp in length
* 21149 21248: gap of 100 bp
* 21249 21841: contig of 593 bp in length
* 21842 21941: gap of 100 bp
* 21942 22950: contig of 1009 bp in length
* 22951 23050: gap of 100 bp
* 23051 24345: contig of 1295 bp in length
* 24346 24445: gap of 100 bp
* 24446 25663: contig of 1218 bp in length
* 25664 32382: contig of 6619 bp in length
* 32383 32483: gap of 100 bp
* 32483 52046: contig of 19564 bp in length
* 52047 52147: gap of 100 bp
* 52147 121530: contig of 69384 bp in length
* 121531 121630: gap of 100 bp
* 121631 155705: contig of 34075 bp in length
* 155706 155805: gap of 100 bp
* 155806 168354: contig of 12549 bp in length.
* Location/Qualifiers
* 1..168354
* /organism="Mus musculus"
* /mol_type="genomic DNA"
* /db_xref="taxon:10090"

```

TITLE
JOURNAL

COMMENT

FEATURES
source

```

misc_feature      /clones="RP24-21117"
1..20209          /clone_lib="RPCi-24 Male Mouse BAC"
/note="assembly_fragment
clone_end:SP6
vector_side:left"
20210..20309      /estimated_length=100
misc_feature      20310..21148
/note="assembly_fragment"
21149..21248      /estimated_length=100
misc_feature      21249..21841
/note="assembly_fragment"
21842..21941      /estimated_length=100
misc_feature      21942..22950
/note="assembly_fragment"
22951..23050      /estimated_length=100
misc_feature      23051..24345
/note="assembly_fragment"
24346..24445      /estimated_length=100
misc_feature      24446..25663
/note="assembly_fragment"
25664..25763      /estimated_length=100
misc_feature      25764..32382
/note="assembly_fragment"
32383..32482      /estimated_length=100
misc_feature      32483..52046
/note="assembly_fragment"
52047..52146      /estimated_length=100
misc_feature      52147..121530
/note="assembly_fragment"
121531..121630    /estimated_length=100
misc_feature      121631..155705
/note="assembly_fragment"
155706..155805    /estimated_length=100
misc_feature      155806..168354
/note="assembly_fragment
clone_end:T7
vector_side:right"

```

```

ORIGIN
Query Match      83.2%; Score 20.8; DB 14; Length 168354;
Best Local Similarity 91.7%; Pred. No. 61;
Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

Qy 1 AAAAAACCAATGACCAACTGGGCTG 24
|||||

```

```

Db 156059 AAGAAACAAATGACCACTGGGCTG 156082

```

```

RESULT 8
AC142146          170429 bp      DNA      linear      ROD 15-MAY-2004
DEFINITION      Mus musculus BAC clone RP24-74D8 from chromosome 9, complete
sequence.
AC142146
AC142146.3      GI:46240972
KEYWORDS      HTG.
SOURCE      Mus musculus (house mouse)
ORGANISM      Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 170429)

```

```

AUTHORS      Isak,A., Bielicki,L., Haglund,K. and Haakenson,W.
TITLE      The sequence of Mus musculus BAC clone RP24-74D8
JOURNAL      Unpublished (2001)
REFERENCE      2 (bases 1 to 170429)
AUTHORS      McPherson,J.D. and Waterston,R.H.
TITLE      Direct Submission
JOURNAL      Submitted (22-MAR-2003) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
REFERENCE      3 (bases 1 to 170429)
AUTHORS      McPherson,J.D. and Waterston,R.H.
TITLE      Direct Submission
JOURNAL      Submitted (23-MAR-2003) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
REFERENCE      4 (bases 1 to 170429)
AUTHORS      Wilson,R.K.
TITLE      Direct Submission
JOURNAL      Submitted (06-APR-2004) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
REFERENCE      5 (bases 1 to 170429)
AUTHORS      Wilson,R.K.
TITLE      Direct Submission
JOURNAL      Submitted (15-MAY-2004) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
COMMENT      On Apr 6, 2004 this sequence version replaced gi:29164701.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu
Contact: submissions@watson.wustl.edu
----- Summary Statistics
Center project name: M_BB0074D08
-----

```

NOTICE:

This sequence was finished as follows unless otherwise noted:
 all regions were double stranded, sequenced with an alternate
 chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by at least
 one plasmid subclone, fosmid clone or direct clone walk sequence.
 Sequence from the Mouse Genome Sequencing Consortium whole genome
 shotgun may have been used to obtain the consensus sequence; and
 the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. Wes Warren,
 Department of Genetics, Washington University, St. Louis MO. For
 additional information about the map position of this sequence, see
 http://genome.wustl.edu

SOURCE INFORMATION:

The RPCi-24 BAC Library has been constructed by Pieter de Jong and
 coworkers (http://www.chori.org) from male C57BL/6J mouse spleen
 and/or brain genomic DNA. The clone and detailed information can be
 obtained from Pieter de Jong and coworkers at http://www.chori.org

NEIGHBORING SEQUENCE INFORMATION:

This sequence is the entire insert of the clone.

```

FEATURES
source
1..170429
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="9"
/map="9"
/clone="RP24-74D8"
/clone_lib="RPCi-24"
1..51_
/rpt_family="L1"
1003..1184
/rpt_family="B2"
2454..2754
repeat_region
repeat_region
repeat_region

```

```
repeat_region /rpt_family="L1"
9564..9667
/rpt_family="B4"
9694..9985
/rpt_family="L1"
10125..10335
/rpt_family="L1"
11552..11666
/rpt_family="B4"
12289..12334
/rpt_family="L1"
13538..13598
/rpt_family="ERV1"
16456..16593
/rpt_family="B4"
16748..17136
/rpt_family="ERVK"
17590..17758
/rpt_family="L1"
18022..18100
/rpt_family="ERVK"
18400..18525
/rpt_family="RMER15"
18527..18816
/rpt_family="L1"
18828..19159
/notes="Unresolved simple sequence repeat."
repeat_region 19147..19464
/rpt_family="RMER15"
20919..21384
/rpt_family="MaLR"
21761..21828
/rpt_family="ERV1"
22079..22305
/rpt_family="B4"
22306..22440
/rpt_family="Alu"
22441..22514
/rpt_family="B4"
22560..22646
/rpt_family="ERVK"
23695..23772
/rpt_family="Alu"
24242..24761
/rpt_family="RMER6A"
24936..26057
/rpt_family="L1"
25976..26056
/notes="Unresolved simple sequence repeat."
repeat_region 29547..29893
/rpt_family="MaLR"
30811..30935
/rpt_family="B4"
33293..33607
/rpt_family="MER2_type"
33770..34174
/rpt_family="MaLR"
34354..34450
/rpt_family="L2"
37257..37539
/notes="Unresolved simple sequence repeat."
repeat_region 39204..39862
/rpt_family="L1"
40201..40267
/rpt_family="ID"
41021..41418
/rpt_family="MaLR"
42084..42461
/rpt_family="MaLR"
42766..42907
/rpt_family="MIR"
42923..43109
/rpt_family="B2"

repeat_region 43782..43830
/rpt_family="L2"
44077..44158
/rpt_family="L1"
45787..45980
/rpt_family="B2"
46247..46447
/rpt_family="MaLR"
46604..47002
/rpt_family="MaLR"
47907..48071
/rpt_family="MaLR"
48783..49214
/rpt_family="L1"
54485..54678
/rpt_family="L2"
56157..56351
/rpt_family="L1"
56444..56688
/rpt_family="L2"
58240..58395
/rpt_family="MaLR"
58416..58578
/rpt_family="B2"
58648..59082
/rpt_family="L1"
65037..65180
/rpt_family="Alu"
65593..65943
/rpt_family="MaLR"
66012..66310
/rpt_family="MaLR"
66722..67335
/rpt_family="MaLR"
67336..67681
/rpt_family="MaLR"
67694..67825
/rpt_family="Alu"
68855..69037
/rpt_family="B2"
68890..69046
/rpt_family="B4"

Query Match 83.2%; Score 20.8; DB 9; Length 170429;
Best Local Similarity 91.7%; Pred.No.61;
Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAACAATGACCACTGGGCTG 24
||| ||||| ||||| |||||
Db 19253 AAGAAACAATGACCAGCTGGGCTG 19276

RESULT 9
AL590454 171970 bp DNA linear HTG 10-JUL-2001
LOCUS Homo sapiens chromosome 1 clone RP11-266010, 15 unordered pieces.
DEFINITION
ACCESSION AL590454
VERSION AL590454.9 GI:14270736
KEYWORDS HTG; HTGS_PHASE1; HTGS_CANCELLED.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE
1 Mclay,K.
AUTHORS Direct Submission
TITLE Submitted (09-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,
JOURNAL CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
COMMENT On May 31, 2001 this sequence version replaced gi:14268249.
----- Genome Center
Center: Sanger Centre
```

Center code: SC
 Web site: <http://www.sanger.ac.uk>
 Contact: hunquary@sanger.ac.uk
 ----- Project Information
 Center project name: BA266010
 ----- Summary Statistics
 Sequencing program: XGAP4; version 4.5
 Sequencing vector: plasmid; L08752; 100% of reads
 Chemistry: Dye-terminator ET-amersham; 3% of reads Chemistry:
 Dye-terminator Big Dye; 96% of reads
 Consensus quality: 168624 bases at least Q40
 Consensus quality: 169393 bases at least Q30
 Consensus quality: 170059 bases at least Q20
 Insert size: 170570; sum-of-contigs
 Insert size: 187521; 2.8% error; agarose-fp
 Quality coverage: 8.38x in Q20 bases; sum-of-contigs Quality
 coverage: 7.68x in Q20 bases; agarose-fp

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 15 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 *
 * 1 6149: contig of 6149 bp in length
 * 6150 6249: gap of 100 bp
 * 6250 12238: contig of 5989 bp in length
 * 12239 12338: gap of 100 bp
 * 12339 22361: contig of 10023 bp in length
 * 22362 22461: gap of 100 bp
 * 22462 35419: contig of 12958 bp in length
 * 35420 35519: gap of 100 bp
 * 35520 41347: contig of 5828 bp in length
 * 41348 41447: gap of 100 bp
 * 41448 44902: contig of 3455 bp in length
 * 44903 45002: gap of 100 bp
 * 45003 68730: contig of 23728 bp in length
 * 68731 68830: gap of 100 bp
 * 68831 81218: contig of 12388 bp in length
 * 81219 81318: gap of 100 bp
 * 81319 104441: contig of 23123 bp in length
 * 104442 104541: gap of 100 bp
 * 104542 109173: contig of 4631 bp in length
 * 109173 109273: gap of 100 bp
 * 109273 118630: contig of 9358 bp in length
 * 118631 118730: gap of 100 bp
 * 118731 137728: contig of 18998 bp in length
 * 137729 137828: gap of 100 bp
 * 137829 145693: contig of 7864 bp in length
 * 145693 145793: gap of 100 bp
 * 145793 154640: contig of 8848 bp in length
 * 154641 154740: gap of 100 bp
 * 154741 171970: contig of 17230 bp in length.
 *
 * Location/Qualifiers
 * 1. 171970
 * /organism="Homo sapiens"
 * /mol_type="genomic DNA"
 * /db_xref="taxon:9606"
 * /chromosome="1"
 * /clone_lib="RPC1-11.1"
 * /clone="RP11-266010"
 * 1. 6149
 * /note="assembly fragment:03712
 * fragment chain:1
 * clone_end:SP6
 * vector_side:left"
 * 6250. 12238
 * /note="assembly fragment:02922
 * fragment chain:1"
 * 12339. 22361
 * /note="assembly fragment:01816

fragment_chain:1"
 22462. 35419
 /note="assembly fragment:00457
 fragment_chain:2"
 35520. 41347
 /note="assembly fragment:03865
 fragment_chain:2"
 41448. 44902
 /note="assembly fragment:02285
 fragment_chain:3"
 45003. 68730
 /note="assembly fragment:03722
 fragment_chain:3"
 68831. 81218
 /note="assembly fragment:01982
 fragment_chain:3"
 81319. 104441
 /note="assembly fragment:02876
 fragment_chain:3"
 104542. 109172
 /note="assembly fragment:03156
 fragment_chain:3"
 109273. 118630
 /note="assembly fragment:00850
 fragment_chain:3"
 118731. 137728
 /note="assembly fragment:02424
 fragment_chain:3"
 137829. 145692
 /note="assembly fragment:04254
 fragment_chain:3"
 145793. 154640
 /note="assembly fragment:03799
 fragment_chain:3"
 154741. 171970
 /note="assembly fragment:00365
 fragment_chain:3"
 clone_end:T7
 vector_side:right"
 ORIGIN
 Query Match 83.2%; Score 20.8; DB 14; Length 171970;
 Best Local Similarity 91.7%; Pred. No. 61;
 Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 AAAAAACATGACCAACTGGCTG 24
 ||||| ||||| ||||| ||||| |||||
 Db 4799 AAAAAAATGACCAACTGGCTG 4822
 RESULT 10
 AC025988 177654 bp DNA linear HTG 10-SEP-2000
 LOCUS Homo sapiens chromosome 1 clone RP11-270L22 map 1, WORKING DRAFT
 DEFINITION
 AC025988
 AC025988.3 GI:10047868
 HTG; HTGS_PHASE1; HTGS_DRAFT.
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homiidae; Homo.
 1. (bases 1 to 177654)
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
 Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
 Boguslavsky,L., Boukhgalter,B., Brown,A., Burkett,G.,
 Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
 Collymore,A., Cooke,P., DeArrellano,K., Dewar,K., Diaz,J.S.,
 REFERENCES
 AUTHORS
 TITLE
 JOURNAL
 REFERENCE
 AUTHORS

Dodge, S., Domino, M., Doyle, M., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L., Grand-Pierre, N., Grant, G., Hagos, B., Heatford, A., Horton, L., Howland, J. C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J., LaRocque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, C., Liu, G., Locke, K., Macdonald, P., Marquis, N., McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPheeters, R., Meldrim, J., Meneus, L., Mihova, T., Miranda, C., Mienga, V., Morrow, J., Murphy, T., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, T. M., Oliver, J., Peterson, K., Pierre, N., Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zimmer, A. and Zody, M.

Direct Submission
Submitted (19-MAR-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Sep 10, 2000 this sequence version replaced gi:7387390.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L7993
Center clone name: 270_L_22

----- Summary Statistics
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 170576 bases at least Q40
Consensus quality: 174464 bases at least Q30
Consensus quality: 176025 bases at least Q20
Insert size: 173000; agarose-fp
Insert size: 176854; sum-of-contigs
Quality coverage: 5.0 in Q20 bases; agarose-fp
Quality coverage: 4.9 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 9 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
* as soon as it is available and the accession number will
* be preserved.

* 1 4177: contig of 4177 bp in length
* 4178 4277: gap of 100 bp
* 4278 52907: contig of 48630 bp in length
* 52908 53007: gap of 100 bp
* 53008 58929: contig of 5922 bp in length
* 58930 59029: gap of 100 bp
* 59030 72333: contig of 13304 bp in length
* 72334 72434: gap of 100 bp
* 72434 85952: contig of 13519 bp in length
* 85953 86052: gap of 100 bp
* 86053 106886: contig of 20834 bp in length
* 106887 106987: gap of 100 bp
* 106987 132487: contig of 25501 bp in length
* 132488 132587: gap of 100 bp
* 132588 175174: contig of 42587 bp in length
* 175175 175275: gap of 100 bp
* 175275 177654: contig of 2380 bp in length.

Location/Qualifiers
1. .177654
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

FEATURES
source

/chromosome="1"
/map="1"

/clone="RP11-270L22"
/clone_lib="RPC1-11 Human Male BAC"

misc_feature

1. .4177
/note="assembly_fragment
clone end:SP6
vector side:left"

gap

4178. .4277
/estimated_length=100

misc_feature

4278. .52907
/note="assembly_fragment"

gap

52908. .53007
/estimated_length=100

misc_feature

53008. .58929
/note="assembly_fragment"

gap

58930. .59029
/estimated_length=100

misc_feature

59030. .72333
/note="assembly_fragment"

gap

72334. .72433
/estimated_length=100

misc_feature

72434. .85952
/note="assembly_fragment"

gap

85953. .86052
/estimated_length=100

misc_feature

86053. .106886
/note="assembly_fragment"

gap

106887. .106986
/estimated_length=100

misc_feature

106987. .132487
/note="assembly_fragment"

gap

132488. .132587
/estimated_length=100

misc_feature

132588. .175174
/note="assembly_fragment"

gap

175175. .175274
/estimated_length=100

misc_feature

175275. .177654
/note="assembly_fragment
clone end:T7
vector side:right"

ORIGIN

Query Match 83.2%; Score 20.8; DB 14; Length 177654;
Best Local Similarity 91.7%; Pred. No. 61;
Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAACAATGACCAACTGGGCTG 24
|||||

Db 92762 AAAAAACAATGACCACTGGGCTG 92785
|||||

RESULT 11

HS75N14/c

LOCUS

HS75N14

DEFINITION

Accession

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

183861 bp DNA linear HTG 10-JUL-2001
Homo sapiens chromosome 12 clone RP1-75N14, 4 unordered pieces.
297199
GI:10045120
HTG; HTGS PHASE1; HTGS_CANCELLED.
Homo sapiens (human)
Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1
Ainscough, R.
Direct Submission
Submitted (08-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Sep 9, 2000 this sequence version replaced gi:9715806.
----- Genome Center

Tel: 6172521477

Fax: 6172580903

Email: kersli@genome.wi.mit.edu

Primer A: No sequence submitted

Primer B: No sequence submitted

STS size: 601

Protocol:

WGS-discovery (WGS):

Paired-end and low-coverage whole genome shotgun reads were generated from 9 breeds (German Shepherd, Rottweiler, Bedlington Terrier, Beagle, Labrador Retriever, English Shepherd, Italian Greyhound, Alaskan Malamute and the Portuguese Water Dog -100,000 each) and five other canids (Chinese, Alaskan, Indian and Spanish Gray Wolf as well as the Californian Coyote).

The WGS reads were placed uniquely on the CanFam1.0 boxer assembly and SNP detection was carried out by SSAHA-SNP. 863872 reads were annotated as STSS and 45941 SNPs were annotated with alleles from the boxer and the breed or canid from which the particular read came. The validation rate for these SNPs was estimated at approximately 98%.

WGA-discovery (WGA) of Boxer/Poodle SNPs: A second set of SNPs was generated using a similar methodology from the 1.5x poodle assembly (Kirkness 2003) were used instead of WGS reads. Since this sequence lacked base quality scores, arbitrary quality scores of phred 40 were assigned before the poodle sequence was placed uniquely on the CanFam1.0 boxer assembly and SNP detection was carried out by SSAHA-SNP. 1637780 SNPs were annotated with alleles from the boxer and the poodle. The validation rate for these SNPs was estimated at approximately TBD%.

Internal-WGA-discovery (I-WGA): A third set of SNPs were discovered by comparing reads in the WGA assembly. SNPs were defined as mismatch positions that had a base quality of ≥ 30 on both reads in a region that aligned without gaps, and with at most one additional mismatch in the ten flanking bases. For each allele, at least one additional read had to confirm it. 731476 SNPs were annotated with alleles between the two boxer alleles. The validation rate for these SNPs was estimated at approximately TBD%.

FEATURES

source

```
1. .601
  /location/Qualifiers
    /organism="Canis familiaris"
    /mol_type="genomic DNA"
    /strain="Beagle"
    /db_xref="taxon:9615"
    /map="9 8-534 44672481-44671955"
    /clone_lib="Beagle"
    <1. .>601
```

ORIGIN

STS

Query Match 80.8%; Score 20.2; DB 10; Length 601;
 Best Local Similarity 88.0%; Pred. No. 1.3e+02;
 Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AAAAAAATGACCACTGGGCTGT 25

Db 201 AAAAAACAGTGACCACTGGGCAGT 177

RESULT 14

AC165059/c

LOCUS

AC165059

DNA

linear

HTG 02-JUL-2005

RESULT 15

AC100243

LOCUS

DEFINITION

AC100243

ACCESSION

VERSION

AC100243

Mus musculus clone RP23-64K4, LOW-PASS SEQUENCE SAMPLING.

AC100243

AC100243.1

GI:17047609

5175 bp

DNA

linear

HTG 22-NOV-2001

DEFINITION

ACCESSION

AC165059

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Phakopsora pachyrhizi clone JG1AFNA-1965M23, WORKING DRAFT SEQUENCE, 2 unordered pieces.

AC165059, 1 GI:68533343

HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.

Phakopsora pachyrhizi

Phakopsora pachyrhizi

Eukaryota; Fungi; Basidiomycota; Urediniomycetes;

Urediniomycetidae; Uredinales; Phakopsoraceae; Phakopsora.

1 (bases 1 to 40745)

DOE Joint Genome Institute.

Unpublished

2 (bases 1 to 40745)

DOE Joint Genome Institute.

Direct Submission

Submitted (02-JUL-2005) Production Genomics Facility, DOE Joint

Genome Institute, 2800 Mitchell Drive B100, Walnut Creek, CA

94598-1698, USA

-----Genome Center

Center: Joint Genome Institute

Center Code: JGI

Web site: http://www.jgi.doe.gov

Project Information

Center Project Name: 4000932

Center clone name: JGI-AFNA_1965M23

Summary Statistics

Consensus quality: 40507 bases at least Q40

Consensus quality: 40595 bases at least Q30

Consensus quality: 40641 bases at least Q20

Estimated insert size: 46000; pulse field gel estimation

Estimated insert size: 40645; sum-of-contigs estimation

Quality coverage: 17.22 in Q20 bases; pulse field gel estimation

Quality coverage: 19.49 in Q20 bases; sum-of-contigs estimation.

* NOTE: This is a 'working draft' sequence. It currently

* consists of 2 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

* as soon as it is available and the accession number will

* be preserved.

* 1 1463: contig of 1463 bp in length

* 1464 1563: gap of unknown length

* 1564 40745: contig of 39182 bp in length.

FEATURES

source

```
1. .40745
  /location/Qualifiers
    /organism="Phakopsora pachyrhizi"
    /mol_type="genomic DNA"
    /db_xref="taxon:170000"
    /clone="JG1AFNA-1965M23"
    /clone_lib="JGI Fommid library AFNA"
    1464. .1563
    /estimated_length=unknown
```

ORIGIN

gap

Query Match 80.8%; Score 20.2; DB 14; Length 40745;
 Best Local Similarity 88.0%; Pred. No. 1.2e+02;
 Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AAAAAAATGACCACTGGGCTGT 25

Db 36386 AAAAAAATGACCACTGGGTTT 36362

```

KEYWORDS      HTG; HTGS PHASE0.
SOURCE         Mus musculus (house mouse)
ORGANISM       Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Murinae; Mus.
REFERENCE      1 (bases 1 to 57175)
AUTHORS        Birren, B., Linton, L., Nuebaum, C. and Lander, E.
TITLE          Mus musculus, clone RP23-64K4
REFERENCE      2 (bases 1 to 57175)
AUTHORS        Birren, B., Linton, L., Nuebaum, C., Lander, E., Ali, A., Allen, N.,
Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Boukhgalter, B.,
Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B.,
Choepeil, Y., Colangelo, M., Collins, S., Collamore, A., Cook, A.,
Cooke, P., DeArellano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S.,
Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S.,
Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
Hagos, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,
Jones, C., Kamat, A., Karatas, A., Kells, C., LaRocque, K.,
Lamazares, R., Landers, T., Lenoczky, J., Levine, R., Liu, G.,
MacLean, C., MacDonald, P., Major, J., Marquis, N., Matthews, C.,
McCarthy, M., McEwan, P., McKernan, K., McPheeters, R., Meldrim, J.,
Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C.,
Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neil, D.,
Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,
Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R.,
Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,
Viel, R., Vo, A., Wilson, N., Wilton, D., Wu, X., Wyman, D., Ye, W. J., Young, G.,
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (22-NOV-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L14443
Center clone name: 64_K_4
-----
* NOTE: This record contains 71 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
*
* 1       709: contig of 709 bp in length
*
* 810     809: gap of 100 bp
*
* 1521    1521: contig of 712 bp in length
*
* 1622    1621: gap of 100 bp
*
* 2345    2345: contig of 724 bp in length
*
* 2446    2445: gap of 100 bp
*
* 3149    3149: contig of 704 bp in length
*
* 3249    3249: gap of 100 bp
*
* 3954    3954: contig of 705 bp in length
*
* 3955    4054: gap of 100 bp
*
* 4055    4768: contig of 714 bp in length
*
* 4769    4868: gap of 100 bp
*
* 4869    5577: contig of 709 bp in length
*
* 5578    5677: gap of 100 bp
*
*
* 5678    6391: contig of 714 bp in length
*
* 6392    6491: gap of 100 bp
*
* 6492    7198: contig of 707 bp in length
*
* 7199    7298: gap of 100 bp
*
* 7299    8014: contig of 716 bp in length
*
* 8015    8115: gap of 100 bp
*
* 8115    8832: contig of 718 bp in length
*
* 8833    8932: gap of 100 bp
*
* 8933    9648: contig of 716 bp in length
*
* 9649    9749: gap of 100 bp
*
* 9749    10452: contig of 704 bp in length
*
* 10453    10552: gap of 100 bp
*
* 10553    11250: contig of 698 bp in length
*
* 11251    11350: gap of 100 bp
*
* 11351    12039: contig of 689 bp in length
*
* 12040    12139: gap of 100 bp
*
* 12140    12851: contig of 712 bp in length
*
* 12852    12951: gap of 100 bp
*
* 12952    13666: contig of 715 bp in length
*
* 13667    13766: gap of 100 bp
*
* 13767    14463: contig of 697 bp in length
*
* 14464    14563: gap of 100 bp
*
* 14564    15278: contig of 715 bp in length
*
* 15279    15378: gap of 100 bp
*
* 15379    16100: contig of 722 bp in length
*
* 16101    16200: gap of 100 bp
*
* 16201    16920: contig of 720 bp in length
*
* 16921    17020: gap of 100 bp
*
* 17021    17119: contig of 699 bp in length
*
* 17120    17819: gap of 100 bp
*
* 17820    18523: contig of 704 bp in length
*
* 18524    18623: gap of 100 bp
*
* 18624    19335: contig of 712 bp in length
*
* 19336    19435: gap of 100 bp
*
* 19436    20251: contig of 716 bp in length
*
* 20252    20361: contig of 710 bp in length
*
* 20362    21061: gap of 100 bp
*
* 21062    21776: contig of 715 bp in length
*
* 21777    21876: gap of 100 bp
*
* 21877    22600: contig of 724 bp in length
*
* 22601    22700: gap of 100 bp
*
* 22701    23421: contig of 721 bp in length
*
* 23422    23521: gap of 100 bp
*
* 23522    24328: contig of 707 bp in length
*
* 24329    24328: gap of 100 bp
*
* 24329    25020: contig of 692 bp in length
*
* 25021    25120: gap of 100 bp
*
* 25121    25810: contig of 690 bp in length
*
* 25811    25910: gap of 100 bp
*
* 25911    26617: contig of 707 bp in length
*
* 26618    26717: gap of 100 bp
*
* 26718    27426: contig of 709 bp in length
*
* 27427    27526: gap of 100 bp
*
* 27527    28243: contig of 717 bp in length
*
* 28244    28343: gap of 100 bp
*
* 28344    29019: contig of 676 bp in length
*
* 29020    29119: gap of 100 bp
*
* 29120    29816: contig of 697 bp in length
*
* 29817    29916: gap of 100 bp
*
* 29917    30625: contig of 709 bp in length
*
* 30626    30725: gap of 100 bp
*
* 30726    31437: contig of 711 bp in length
*
* 31437    31536: gap of 100 bp
*
* 31537    32241: contig of 705 bp in length
*
* 32242    32341: gap of 100 bp
*
* 32342    33056: contig of 715 bp in length
*
* 33057    33156: gap of 100 bp
*
* 33157    33864: contig of 708 bp in length
*
* 33865    33964: gap of 100 bp
*
* 33965    34671: contig of 707 bp in length
*
* 34672    34771: gap of 100 bp
*
* 34772    35481: contig of 710 bp in length

```

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS

```

35482 35581: gap of 100 bp
35582: contig of 705 bp in length
36287 36386: gap of 100 bp
36387 37086: contig of 700 bp in length
37087 37186: gap of 100 bp
37187 37899: contig of 713 bp in length
37900 37999: gap of 100 bp
38000 38713: contig of 714 bp in length
38714 38813: gap of 100 bp
38814 39505: contig of 692 bp in length
39506 39606: gap of 100 bp
39607 40306: contig of 701 bp in length
40307 40406: gap of 100 bp
40407 41112: contig of 706 bp in length
41113 41212: gap of 100 bp
41213 41917: contig of 705 bp in length
41918 42017: gap of 100 bp
42019 42724: contig of 707 bp in length
42725 42824: gap of 100 bp
42825 43539: contig of 715 bp in length
43540 43639: gap of 100 bp
43640 44354: contig of 715 bp in length
44355 44454: gap of 100 bp
44455 45135: contig of 681 bp in length
45136 45235: gap of 100 bp
45236 45947: contig of 712 bp in length
45948 46047: gap of 100 bp
46048 46748: contig of 701 bp in length
46749 46848: gap of 100 bp
46849 47552: contig of 704 bp in length
47553 47653: gap of 100 bp
47653 48351: contig of 699 bp in length
48352 48451: gap of 100 bp
48452 49122: contig of 671 bp in length
49123 49222: gap of 100 bp
49223 49935: contig of 713 bp in length
49936 50035: gap of 100 bp
50035 50747: contig of 712 bp in length
50748 50847: gap of 100 bp
50848 51520: contig of 673 bp in length
51521 51621: gap of 100 bp
51621 52336: contig of 716 bp in length
52337 52436: gap of 100 bp
52437 53137: contig of 701 bp in length
53138 53237: gap of 100 bp
53238 53953: contig of 716 bp in length
53954 54053: gap of 100 bp
54054 54756: contig of 703 bp in length
54757 54856: gap of 100 bp
54857 55563: contig of 707 bp in length

Query Match      80.8%; Score 20.2; DB 14; Length 57175;
Best Local Similarity 88.0%; Pred. No. 1.2e+02;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AAAAAACAATGACCACTGGCTGT 25
Db 20553 AAGAAACAAGACCACTGAGCTGT 20577

```

```

RESULT 16
AC111185
LOCUS Homo sapiens chromosome 17 clone RP11-1109H12 map 17, LOW-PASS
DEFINITION AC111185 62356 bp DNA linear HTG 18-FEB-2002
ACCESSION AC111185
VERSION AC111185.1 GI:186999950
KEYWORDS HTG; HTGS_PHASE0.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.

```

```

1 (bases 1 to 62356)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens chromosome 17, clone RP11-1109H12
Unpublished
2 (bases 1 to 62356)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhgalter,B.,
Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
Choepel,Y., Collangelo,M., Collins,S., Collymore,A., Cook,A.,
Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., Lakocque,K., Lamazares,R.,
Landers,T., Lehoczy,J., Levine,R., Liu,G., Maclean,C.,
Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M.,
McEwan,P., McKernan,K., Meldrum,J., Meneus,L., Mihova,T.,
Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J.,
Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C.,
Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,
Rosetti,M., Roy,A., Santos,R., Schauer,S., Schuback,R., Seaman,S.,
Severy,P., Spencer,B., Strange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigglio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (18-FEB-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L25448
Center clone name: 1109_H_12
-----
* NOTE: This record contains 79 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
* 1 699: contig of 699 bp in length
* 700 799: gap of 100 bp
* 800 1497: contig of 698 bp in length
* 1498 1597: gap of 100 bp
* 1598 2300: contig of 703 bp in length
* 2301 2400: gap of 100 bp
* 2401 3085: contig of 685 bp in length
* 3086 3185: gap of 100 bp
* 3186 3868: contig of 683 bp in length
* 3869 3969: gap of 100 bp
* 3969 4652: contig of 684 bp in length
* 4653 4753: gap of 100 bp
* 4753 5461: contig of 709 bp in length
* 5462 5562: gap of 100 bp
* 5562 6272: contig of 710 bp in length
* 6272 7067: gap of 100 bp
* 7068 7168: contig of 696 bp in length
* 7168 7852: contig of 100 bp
* 7852 7952: contig of 685 bp in length
* 7952 7952: gap of 100 bp

```

TITLE
JOURNAL
COMMENT

```
* 7953 8667: contig of 715 bp in length
* 8668 8767: gap of 100 bp
* 9471: contig of 704 bp in length
* 9472 9571: gap of 100 bp
* 9572 10278: contig of 707 bp in length
* 10279 11060: contig of 682 bp in length
* 11061 11160: gap of 100 bp
* 11161 11856: contig of 696 bp in length
* 11857 11956: gap of 100 bp
* 11957 12635: contig of 679 bp in length
* 12636 12735: gap of 100 bp
* 12736 13444: contig of 709 bp in length
* 13445 13544: gap of 100 bp
* 13545 14240: contig of 696 bp in length
* 14241 14340: gap of 100 bp
* 14341 15010: contig of 670 bp in length
* 15011 15110: gap of 100 bp
* 15111 15799: contig of 689 bp in length
* 15800 15899: gap of 100 bp
* 15900 16562: contig of 663 bp in length
* 16563 16662: gap of 100 bp
* 16664 17358: contig of 696 bp in length
* 17359 17458: gap of 100 bp
* 17459 18164: contig of 706 bp in length
* 18165 18264: gap of 100 bp
* 18265 18945: contig of 681 bp in length
* 18946 19045: gap of 100 bp
* 19046 19726: contig of 681 bp in length
* 19727 19826: gap of 100 bp
* 19827 20515: contig of 689 bp in length
* 20516 20615: gap of 100 bp
* 20616 21326: contig of 711 bp in length
* 21327 21426: gap of 100 bp
* 21427 22139: contig of 713 bp in length
* 22140 22239: gap of 100 bp
* 22240 22925: contig of 686 bp in length
* 22926 23025: gap of 100 bp
* 23026 23703: contig of 678 bp in length
* 23704 23803: gap of 100 bp
* 23804 24507: contig of 704 bp in length
* 24508 24607: gap of 100 bp
* 24608 25308: contig of 701 bp in length
* 25309 25408: gap of 100 bp
* 25409 26075: contig of 667 bp in length
* 26076 26175: gap of 100 bp
* 26176 26863: contig of 688 bp in length
* 26864 26963: gap of 100 bp
* 26964 27652: contig of 689 bp in length
* 27653 27753: gap of 100 bp
* 27753 28466: contig of 714 bp in length
* 28467 28566: gap of 100 bp
* 28567 29273: contig of 707 bp in length
* 29274 29373: gap of 100 bp
* 29374 30050: contig of 677 bp in length
* 30051 30150: gap of 100 bp
* 30151 30831: contig of 681 bp in length
* 30832 30931: gap of 100 bp
* 30932 31592: contig of 661 bp in length
* 31593 31692: gap of 100 bp
* 31693 32381: contig of 689 bp in length
* 32382 32481: gap of 100 bp
* 32482 33189: contig of 708 bp in length
* 33190 33289: gap of 100 bp
* 33290 33983: contig of 694 bp in length
* 33984 34083: gap of 100 bp
* 34084 34773: contig of 690 bp in length
* 34774 34873: gap of 100 bp
* 34874 35551: contig of 678 bp in length
* 35552 36351: gap of 100 bp
* 36352 36552: contig of 707 bp in length
* 36353 36459: gap of 100 bp
* 36459 37150: contig of 692 bp in length
```

```
* 37151 37250: gap of 100 bp
* 37251 37927: contig of 677 bp in length
* 37928 38027: gap of 100 bp
* 38028 38685: contig of 658 bp in length
* 38686 38785: gap of 100 bp
* 38786 39489: contig of 704 bp in length
* 39490 39589: gap of 100 bp
* 39590 40287: contig of 698 bp in length
* 40288 40387: gap of 100 bp
* 40388 41080: contig of 693 bp in length
* 41081 41180: gap of 100 bp
* 41181 41863: contig of 683 bp in length
* 41864 41963: gap of 100 bp
* 41964 42642: contig of 679 bp in length
* 42643 42742: gap of 100 bp
* 42743 43415: contig of 673 bp in length
* 43416 43515: gap of 100 bp
* 43516 44199: contig of 684 bp in length
* 44200 44299: gap of 100 bp
* 44300 44993: contig of 694 bp in length
* 44994 45093: gap of 100 bp
* 45094 45785: contig of 692 bp in length
* 45786 45885: gap of 100 bp
* 45886 46565: contig of 680 bp in length
* 46566 47360: contig of 695 bp in length
* 47361 47460: gap of 100 bp
* 47461 48140: contig of 680 bp in length
* 48141 48240: gap of 100 bp
* 48241 48922: contig of 682 bp in length
* 48923 49022: gap of 100 bp
* 49023 49711: contig of 689 bp in length
* 49712 49811: gap of 100 bp
* 49812 50496: contig of 685 bp in length
* 50497 50596: gap of 100 bp
* 50597 51294: contig of 698 bp in length
* 51295 51394: gap of 100 bp
* 51395 52092: contig of 698 bp in length
* 52093 52192: gap of 100 bp
* 52193 52905: contig of 713 bp in length
* 52906 53005: gap of 100 bp
* 53006 53703: contig of 698 bp in length
* 53704 53803: gap of 100 bp

Query Match      80.8%; Score 20.2; DB 14; Length 62356;
Best Local Similarity 88.0%; Pred. No. 1.2e+02;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAACATGACCAACTGGGCTGT 25
    ||||| ||||| ||||| ||||| |||||
Db 21910 AAAAAACATGACCAACTGGGCTGT 21934

RESULT 17
AL356141/c
LOCUS
DEFINITION
  Human DNA sequence from clone Rp11-269L3 on chromosome 10, complete
  sequence.
ACCESSION
  AL356141
VERSION
  AL356141.9 GI:15617255
KEYWORDS
  HTG.
SOURCE
  Homo sapiens (human)
  ORGANISM
    Homo sapiens
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
    Hominiidae; Homo.
REFERENCE
  1 (bases 1 to 87871)
AUTHORS
  Heath, P.
TITLE
  Direct Submission
JOURNAL
  Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
  Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vega@sanger.ac.uk
  Clone requests: clonerequest@sanger.ac.uk
COMMENT
  On Sep 13, 2001 this sequence version replaced gi:14970334.
```

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:
 Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at
<http://www.sanger.ac.uk/projects/C-elegans/wormpep> This sequence was generated from part of bacterial clone contigs of human chromosome 10, constructed by the Sanger Centre Chromosome 10 Mapping Group. Further information can be found at
<http://www.sanger.ac.uk/HGP/Chr10>
 RP11-269L3 is from the library RP11-11.1 constructed by the group of Pieter de Jong. For further details see
<http://www.chori.org/bacpac/home.htm>
 VECTOR: pBACe3.6

----- Genome Center
 Center: Wellcome Trust Sanger Institute
 Center code: SC
 Web site: <http://www.sanger.ac.uk>
 Contact: vega@sanger.ac.uk

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

FEATURES

source
 1. 87871
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="10"
 /clone="RP11-269L3"
 /clone_lib="RP11-11.1"
 /note="Clone_left_end: RP11-269L3"
 85872
 /note="Clone_left_end: RP11-47B24"
 ORIGIN
 Query Match 80.8%; Score 20.2; DB 8; Length 87871;
 Best Local Similarity 88.0%; Pred. No. 1.2e+02;
 Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 Qy 1 AAAAAACAATGACCACTGGCTGT 25
 |||||
 Db 6325 AAAAAACAATGACCACTGGCTGT 6301
 |||||

RESULT 18

AC154170
 LOCUS
 DEFINITION Alligator mississippiensis clone VMRC8-371122, WORKING DRAFT
 AC154170 DNA linear HTG 24-FEB-2005
 SEQUENCE, 8 ordered pieces.
 AC154170
 AC154170.2 GI:50223218
 HTG; HTGS PHASE2; HTGS DRAFT.
 KEYWORDS Alligator mississippiensis (American alligator)
 SOURCE Alligator mississippiensis
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Crocodylia; Alligatorinae; Alligator.
 1 (Bases 1 to 134231)
 Antonellis,A., Ayele,K., Benjamin,B., Blakesley,R.W., Boake,A., Bouffard,G.G., Brinkley,C., Brooks,S., Chu,G., Coleman,H., Engle,J., Gestole,M., Guan,X., Gupta,J., Gutierrez,P., Haghighi,P., Han,J., Hansen,N., Ho,S.-L., Hu,P., Hunter,G., Hurle,B., Idol,J.R., Jones,C., Kwong,P., Laric,P., Larson,S., Lee-Lin,S.-Q., Legaspi,R., Madden,M., Maduro,Q.L., Maduro,V.B., Margulies,E.H., Masiello,C., Maakeri,B., McDowell,J., Mullikin,J.C., Oestreicher,J.S., Park,M., Portnoy,M.E., Prasad,A., Puri,O., Reddix-Dugue,N., Rosas,B., Schandler,K., Schueler,M.G., Sison,C., Stantripop,S., Stephen,E., Taye,A., Thomas,J.W., Thomas P.J., Tsipouri,V., Vogt,J.L., Wetherby,K.D., Young,A. and Green,E.D.

TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

NISC Comparative Sequencing Initiative
 Unpublished
 2 (bases 1 to 134231)
 Green,E.D.
 Direct Submission
 Submitted (28-DEC-2004) NIH Intramural Sequencing Center, 5625 Fishers Lane, Rockville, MD 20852, USA
 3 (bases 1 to 134231)
 Green,E.D.
 Direct Submission
 Submitted (24-FEB-2005) NIH Intramural Sequencing Center, 5625 Fishers Lane, Rockville, MD 20852, USA
 On Feb 24, 2005 this sequence version replaced gi:56806706.
 ----- Genome Center
 Center: NIH Intramural Sequencing Center
 Center code: NISC
 Web site: <http://www.nisc.nih.gov>
 Contact: nisc.zoo@nhgri.nih.gov
 ----- Project Information
 Center project name: gnt
 Center clone name: 371122

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig, has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is generally based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics

Sequencing vector: plasmid; n/a; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 133035 bases at least Q40
 Consensus quality: 133354 bases at least Q30
 Consensus quality: 133486 bases at least Q20
 Insert size: 152000; agarose-ep
 Insert size: 133531; sum-of-contigs
 Quality coverage: 10.18x in Q20 bases; agarose-ep
 Quality coverage: 11.59x in Q20 bases; sum-of-contigs

----- NOTE: This is a 'working draft' sequence. It currently
 * consists of 8 contigs. Gaps between the contigs
 * are represented as runs of N. The order of the pieces
 * is believed to be correct as given, however the sizes
 * of the gaps between them are based on estimates that have
 * provided by the submittor.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.
 1 26879: contig of 26879 bp in length
 26880 26979: gap of unknown length
 26980 29523: contig of 2544 bp in length
 29524 29623: gap of unknown length
 29624 67124: contig of 37401 bp in length
 67025 67124: gap of unknown length
 67125 92702: contig of 25578 bp in length
 92703 92802: gap of unknown length
 92803 95901: contig of 3099 bp in length
 95902 96001: gap of unknown length
 96002 100291: contig of 4290 bp in length
 100292 100391: gap of unknown length
 100392 129050: contig of 28659 bp in length
 129051 129151: gap of unknown length
 129151 134231: contig of 5081 bp in length.
 Location/Qualifiers
 1. 134231
 /organisms="Alligator mississippiensis"

FEATURES
 source

```

/mol_type="genomic DNA"
/db_xref="taxon:8496"
/clone="VMRC8-371122"
/clone_lib="VMRC8"
1. .35894
/note="BAC resource: http://bacpac.chori.org/"
misc_feature
1. .35894
/note="clone overlaps with GenBank Accession Number
AC151169 clone VMRC8-246114 (center project name gns)"
misc_feature
1. .26879
/note="assembly_fragment
vector_side:left"
clone_end:77
26880..26979
/estimated_length=unknown
misc_feature
26980..29523
/note="assembly_fragment"
29524..29623
/estimated_length=unknown
misc_feature
29624..67024
/note="assembly_fragment"
67025..67124
/estimated_length=unknown
misc_feature
67125..92702
/note="assembly_fragment"
92703..92802
/estimated_length=unknown
misc_feature
92803..95901
/note="assembly_fragment"
95902..96001
/estimated_length=unknown
misc_feature
96002..100291
/note="assembly_fragment"
100292..100391
/estimated_length=unknown
misc_feature
100392..129050
/note="assembly_fragment"
129051..129150
/estimated_length=unknown
misc_feature
129151..134231
/note="assembly_fragment
clone_end:SP6
vector_side:right"

ORIGIN
Query Match      80.8%; Score 20.2; DB 14; Length 134231;
Best Local Similarity 88.0%; Pred. No. 1.2e+02;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAAATGACCAACTGGGCTGT 25
Db 52652 AAAAAAGAAAGACCATCTGGGCTGT 52676

RESULT 19
AC090735
LOCUS AC090735 149089 bp DNA linear PRI 20-FEB-2002
DEFINITION Homo sapiens chromosome 8, clone RP11-1145L24, complete sequence.
ACCESSION AC090735
VERSION AC090735.5 GI:18767535
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
REFERENCE
1 (bases 1 to 149089)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Homo sapiens chromosome 8, clone RP11-1145L24
JOURNAL Unpublished
REFERENCE
2 (bases 1 to 149089)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,
Barna,N., Bastien,V., Boguslavskiy,L., Boukhgalter,B., Brown,A.,
Camarata,J., Campopiano,A., Choepel,Y., Colangelo,M., Collins,S.,

```

```

Collamore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J.,
Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Heaford,A., Horton,B., Hulme,W., Iliev,I., Johnson,R.,
Jones,C., Karatas,A., LaRoque,K., Lamazares,R., Landers,T.,
Lehoczy,J., Levine,R., Liu,G., Maclean,C., Macdonald,P.,
Lehoczy,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K.,
McPheeters,R., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H.,
O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R.,
Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M.,
Roy,A., Santos,R., Schauer,S., Schuback,R., Seaman,S., Severy,P.,
Sougez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A.,
Wilson,B., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A.,
Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (09-MAR-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 149089)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Boguslavskiy,L., Boukhgalter,B.,
Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gort,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., LaRoque,K., Lamazares,R.,
Landers,T., Lehoczy,J., Levine,R., Liu,G., Maclean,C.,
Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M.,
McEwan,P., McKernan,K., Meldrim,J., Meneus,L., Mihova,T.,
Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J.,
Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C.,
Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,
Rosetti,M., Roy,A., Santos,R., Schauer,S., Schuback,R., Seaman,S.,
Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (20-FEB-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Feb 20, 2002 this sequence version replaced gi:18653581.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L12821
Center clone name: 1145_L_24
-----
Location/Qualifiers
1. .149089
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="8"
/map="8"
/clone="RP11-1145L24"
/clone_lib="RP11-11 Human Male BAC"
complement(497..603)
/rpt_family="L2"
repeat_region 839..885
repeat_region /rpt_family="L2"

```

FEATURES

source

repeat_region

repeat_region


```

repeat_region 1542..1579
/rpt_family="CT-rich"
complement(1616..1924)
/rpt_family="ALUSC"
repeat_region complement(4056..4126)
/rpt_family="L2"
repeat_region 4128..4191
/rpt_family="GA-rich"
repeat_region 4293..4344
/rpt_family="L1PA6"
complement(4458..4635)
/rpt_family="L1PA10"
repeat_region complement(5395..5571)
/rpt_family="FRAM"
repeat_region 5588..5848
/rpt_family="LTR16C"
complement(6035..6096)
/rpt_family="MIR"
repeat_region 6456..7176
/rpt_family="L2"
repeat_region 7182..7569
/rpt_family="L1MC4a"
complement(7927..8978)
/rpt_family="L1P4"
repeat_region complement(8979..9113)
/rpt_family="FLAM C"
repeat_region complement(9114..9948)
/rpt_family="L1P4"
repeat_region 9949..10342
/rpt_family="L1PA16"
complement(10545..10779)
/rpt_family="L1PA5"
repeat_region complement(13384..13882)
/rpt_family="MLT1D"
repeat_region 14543..15079
/rpt_family="L2"
repeat_region 15694..15714
/rpt_family="AT-rich"
repeat_region 16311..16425
/rpt_family="MIR"
repeat_region 18284..18331
/rpt_family="MIR"
repeat_region 18955..19223
/rpt_family="ALUSx"
complement(19282..19300)
/note="<30 qual SNGL region"
/note="probably T, possibly C"
repeat_region complement(19302..19595)
/rpt_family="AluJb"
repeat_region 19673..19769
/rpt_family="MIR"
repeat_region 20591..20907
/rpt_family="AluJb"
complement(22175..22233)
/rpt_family="L2"
repeat_region 23342..23692
/rpt_family="L1"
repeat_region 23704..24007
/rpt_family="L1MB"
repeat_region 24024..24342
/rpt_family="L1MB"
repeat_region 24360..24561
/rpt_family="L1MB"
repeat_region 24691..25221
/rpt_family="L1MB"
repeat_region complement(26037..26225)
/rpt_family="Charlie8"
repeat_region 26445..26480
/rpt_family="TC)n"
repeat_region 27413..33443
/rpt_family="L1HS"
repeat_region complement(33546..33700)
/rpt_family="MERSA"
complement(35431..35652)
/rpt_family="MIR"
repeat_region complement(35984..36109)
/rpt_family="L2"
repeat_region complement(37679..37730)
/rpt_family="L2"
repeat_region complement(37809..37980)
/rpt_family="MIR"
repeat_region complement(37990..38222)
/rpt_family="AluJb"
complement(38258..38358)
/rpt_family="MIR"
repeat_region complement(38653..39585)
/rpt_family="L1MB1"
repeat_region complement(39586..39661)
/rpt_family="ORSL"
complement(40402..40703)
/rpt_family="AluSc"
unsure 40591..40595
/note="<30 qual SNGL region"
unsure 40634..40716
/note="single clone coverage"
unsure 40655..40660
/note="<30 qual single clone coverage"
unsure 40682..40705
/note="<30 qual single clone coverage"
unsure 40709..40716
/note="<30 qual single clone coverage"
repeat_region 40758..40941
/rpt_family="FRAM"
repeat_region complement(41497..41634)
/rpt_family="FLAM C"
repeat_region complement(41889..42129)
/rpt_family="L2"
repeat_region complement(42956..43280)

Query Match 80.8%; Score 20.2; DB 8; Length 149089;
Best Local Similarity 88.0%; Pred. No. 1.2e+02;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAAATGACCAACTGGCTGT 25
|||||
Db 91563 AAAAAATCAATGACTATCGGCTGT 91587

RESULT 20
AC012339/c
LOCUS AC012339 179818 bp DNA linear PRI 30-DEC-2001
DEFINITION Homo sapiens chromosome 8, clone RP11-31K23, complete sequence.
ACCESSION AC012339
VERSION AC012339.8 GI:17998717
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
1 (bases 1 to 179818)
Birren, B., Linton, L., Nusbaum, C. and Lander, E.
AUTHORS Birren, B., Linton, L., Nusbaum, C. and Lander, E.
TITLE Homo sapiens chromosome 8, clone RP11-31K23
JOURNAL Unpublished
REFERENCE
2 (bases 1 to 179818)
Birren, B., Linton, L., Nusbaum, C., Allen, N., Anderson, M.,
Baldwin, J., Barna, N., Beckerly, R., Boguslavsky, L., Boukhgalter, B.,
Brown, A., Castle, A., Colangelo, R., Collins, S., Collymore, A.,
Cooke, P., DeArellano, K., Dewar, K., Domino, M., Donelan, L., Doyle, M.,
Perreira, P., FitzHugh, W., Forrest, C., Funke, R., Gage, D.,
Galagan, J., Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L.,
Howland, J. C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J.,
Lehoczky, J., Lieu, C., Locke, K., Macdonald, P., Marquis, N.,
McEwan, P., McGurk, A., McKernan, K., McLaughlin, J., Meldrim, J.,
Morrow, J., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P.,

```



```
repeat_region /rpt_family="L3"
29373..29546
repeat_region /rpt_family="MIR"
complement(30916..31155)
repeat_region /rpt_family="MIR"
31657..31932
repeat_region /rpt_family="L2"
31999..32168
repeat_region /rpt_family="L2"
33449..33692
repeat_region /rpt_family="MIR"
complement(34755..35194)
repeat_region /rpt_family="MLT2F"
35805..35842
repeat_region /rpt_family="AT-rich"
38261..38368
repeat_region /rpt_family="LRI16C"
38413..38432
repeat_region /rpt_family="(TG)n"
38601..38626
repeat_region /rpt_family="(TA)n"
38633..39406
repeat_region /rpt_family="LLMC4a"
39407..39699
repeat_region /rpt_family="AluSg"
39700..40505
repeat_region /rpt_family="LLMC4a"
complement(40516..40992)
repeat_region /rpt_family="MLT1D"

Query Match 80.8%; Score 20.2; DB 8; Length 179818;
Best Local Similarity 88.0%; Pred. No. 1.2e+02;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAACAATGACCAACTGGGCTGT 25
||| ||||| ||||| |||||
Db 67424 AACAACTAATGACCAACTGGGCTGT 67400

RESULT 21
AC068709/c 180583 bp DNA linear HTG 22-MAY-2002
LOCUS Homo sapiens chromosome 10 clone RP11-431P18, WORKING DRAFT
DEFINITION SEQUENCE, 4 unordered pieces.
ACCESSION AC068709.4 GI:14290388
VERSION HTG; HTGS PHASE1; HTGS_DRAFT; HTGS_FULLTOP; HTGS_ACTIVEFIN;
KEYWORDS HTGS CANCELLED.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
1 (bases 1 to 180583)
Smith, D.R.
Genome Therapeutics Corporation Sequencing Center: Human Genome
Sequencing Data
Unpublished
2 (bases 1 to 180583)
Smith, D.R.
Direct Submission
Submitted (07-MAY-2000) Genome Therapeutics Corporation, 100 Beaver
Street, Waltham, MA 02453, USA
On Jun 5, 2001 this sequence version replaced gi:14150408.
-----
Center: Genome Therapeutics Corporation
Center code: GTC
Web site: http://www.genomecorp.com/
Contact: gtc-seqcenter@genomecorp.com
-----
Center project name: hg293
-----
Summary Statistics
Sequencing vector: N/A

Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 990315
Consensus quality: 176000 bases at least Q40
Consensus quality: 177396 bases at least Q30
Consensus quality: 178472 bases at least Q20
Insert size: 180333; sum-of-contigs
Quality coverage: 6.1x in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1524: contig of 1524 bp in length
* 1525: gap of unknown length
* 1625: contig of 32742 bp in length
* 34366: gap of unknown length
* 34367: contig of 60856 bp in length
* 95322: gap of unknown length
* 95323: contig of 85161 bp in length.
* 95423: 180583: contig of 85161 bp in length.
*
* Location/Qualifiers
* 1..180583
* /organism="Homo sapiens"
* /mol_type="genomic DNA"
* /db_xref="taxon:9606"
* /chromosome="10"
* /clone_lib="RPC1-11"
* /clone="RP11-431P18"
* 1..1524
* /note="assembly_name:Contig3"
* 1525..1624
* /estimated_length=unknown
* 1625..34366
* /note="assembly_name:Contig4"
* 34367..34466
* /estimated_length=unknown
* 34467..95322
* /note="assembly_name:Contig5"
* 95323..95422
* /estimated_length=unknown
* 95423..180583
* /note="assembly_name:Contig6"

FEATURES
source
misc_feature
gap
misc_feature
gap
misc_feature
gap
misc_feature
ORIGIN
Query Match 80.8%; Score 20.2; DB 14; Length 180583;
Best Local Similarity 88.0%; Pred. No. 1.2e+02;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAACAATGACCAACTGGGCTGT 25
||| ||||| ||||| |||||
Db 65206 AAAAAACAATGACCACTGAGGTGT 65182

RESULT 22
AC013562/c 188755 bp DNA linear PRI 27-AUG-2002
LOCUS Homo sapiens chromosome 8, clone RP11-313C15, complete sequence.
DEFINITION AC013562
ACCESSION AC013562
VERSION AC013562.6 GI:13443208
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
1 (bases 1 to 188755)
Birren, B., Linton, L., Nusbaum, C. and Lander, E.
Homo sapiens chromosome 8, clone RP11-313C15
Unpublished
2 (bases 1 to 188755)
```

AUTHORS

Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M., Baldwin, J., Barna, N., Beckerly, R., Boguslavskiy, L., Boukhgalter, B., Brown, A., Castie, A., Colangelo, M., Collins, S., Collymore, A., Cooke, P., DeArelano, K., Dewar, K., Domino, M., Donelan, L., Doyle, M., Ferreira, P., FitzHugh, W., Forrest, C., Funke, R., Gage, D., Galagan, J., Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L., Howland, J. C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J., Lehoczy, J., Lieu, C., Locke, K., Macdonald, P., Marquis, N., McEwan, P., McKernan, K., McKernan, K., McLaughlin, J., Meldrim, J., Morrow, J., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P., Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Seaman, S., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Tesfaye, S., Tirrell, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X., Wyman, D., Ye, W. J., Zimmer, A. and Zody, M.

TITLE

Direct Submission
 JOURNAL
 Submitted (13-NOV-1999) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE

JOURNAL

3 (bases 1 to 188755)

COMMENT

Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, S., Barna, N., Bastien, V., Boguslavskiy, L., Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P., DeArelano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Karatas, A., LaRocque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G., Maclean, C., Macdonald, P., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., McPheeters, R., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schuback, R., Seaman, S., Severy, P., Sougne, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Straus, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE

Direct Submission
 JOURNAL
 Submitted (01-APR-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE

JOURNAL

4 (bases 1 to 188755)
 Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, S., Barna, N., Bastien, V., Boguslavskiy, L., Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P., DeArelano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Karatas, A., LaRocque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G., Maclean, C., Macdonald, P., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., McPheeters, R., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schuback, R., Seaman, S., Severy, P., Sougne, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Straus, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE

Direct Submission
 JOURNAL
 Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE

JOURNAL

5 (bases 1 to 188755)
 Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, S., Barna, N., Bastien, V., Boguslavskiy, L., Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P., DeArelano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S., Ferreira, P., FitzHugh, W., Gage, D.,

Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Karatas, A., LaRocque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G., Maclean, C., Macdonald, P., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., McPheeters, R., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schuback, R., Seaman, S., Severy, P., Sougne, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Straus, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
 Submitted (27-AUG-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Mar 24, 2001 this sequence version replaced gi:13399376.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L2374
 Center clone name: 313_C_15

----- Location/Qualifiers
 1..188755
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="8"
 /map="8"
 /clone="RP11-313C15"
 /clone_lib="RPC1-11 Human Male BAC"
 117..138
 /rpt_family="AT_rich"
 191..213
 /rpt_family="AT_rich"
 complement(239..866)
 /rpt_family="MER67B"
 complement(1061..1271)
 /rpt_family="AluJo"
 1272..1335
 /rpt_family="AT_rich"
 2094..2293
 /rpt_family="L1PA16"
 complement(3162..3324)
 /rpt_family="MIR3"
 3931..3954
 /rpt_family="AT_rich"
 4511..4829
 /rpt_family="L2"
 complement(4809..4895)
 /rpt_family="MER91B"
 complement(4898..5046)
 /rpt_family="MER91B"
 5047..5327
 /rpt_family="L2"
 complement(5501..5920)
 /rpt_family="MLT1C"
 5974..6265
 /rpt_family="L2"
 6287..6489
 /rpt_family="MER58A"
 complement(6811..6902)
 /rpt_family="MIR"

```

repeat_region 7489. .8170
/rpt_family="L1PA15"
repeat_region complement(8539. .8655)
/rpt_family="L2"
repeat_region 9751. .9786
/rpt_family="AT_rich"
repeat_region 10565. .12681
/rpt_family="Tigger1"
repeat_region complement(13712. .14008)
/rpt_family="AluSx"
repeat_region 14599. .14639
/rpt_family="(TTTTG)n"
repeat_region complement(15742. .15915)
/rpt_family="MIR"
repeat_region complement(15920. .15930)
/rpt_family="L2"
unsure /note="Less than 30 qual SNGL region"
complement(17339. .18108)
/rpt_family="L2"
repeat_region 18277. .18333
/rpt_family="L2"
repeat_region complement(19352. .19478)
/rpt_family="AluSp/q"
repeat_region 19684. .19706
/rpt_family="(TG)n"
repeat_region 19707. .19766
/rpt_family="(GA)n"
repeat_region 19767. .19923
/rpt_family="MIR"
repeat_region 19922. .20636
/rpt_family="L2"
repeat_region complement(20637. .21232)
/rpt_family="MER77"
repeat_region 21233. .21517
/rpt_family="L2"
repeat_region 21617. .21649
/rpt_family="(TG)n"
repeat_region 22651. .22807
/rpt_family="MIR3"
repeat_region 23200. .23379

```

```

Query Match 80.8%; Score 20.2; DB 8; Length 188755;
Best Local Similarity 88.0%; Pred. No. 1.2e+02;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 1 AAAAAACATGACCAACTGGGCTGT 25
||||| ||||| ||||| ||||| |||||
Db 170134 AAAAAACATGACTATCTGGGCTGT 170110

```

```

RESULT 23
AC023376 189963 bp DNA linear HTG 04-JUL-2001
LOCUS Homo sapiens chromosome 8 clone RP11-687M23 map 8, WORKING DRAFT
DEFINITION SEQUENCE, 4 unordered pieces.
AC023376
AC023376 4 GI:14595907
KEYWORDS HTG; HTGS_P1ASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 189963)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Abramam,H., Allen,N.,
TITLE Anderson,S., Baldwin,J., Barna,N., Beda,F., Boguslavskiy,L.,
JOURNAL Boukhgalter,B., Brown,A., Burkett,G., Campopiano,A., Castle,A.,
REFERENCE Choepl,Y., Colangelo,M., Collins,S., Collamore,A., Cooke,P.,
AUTHORS Dearellano,K., Dewar,K., Dodge,S., Domino,M., Doyle,M.,
Fenster,J., Ferreira,P., FitzHugh,W., Forrest,C., Gage,D.,

```

Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L., Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J., Landers,T., Largocque,K., Lehoczy,J., Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N., McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R., Meldrim,J., Meneas,L., Mihova,T., O'Donnell,P., O'Neill,D., Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., Pisan,C., Pollara,V., Raymond,C., Peterson,K., Pierre,N., Rothman,D., Roy,A., Santos,R., Schauer,S., Riley,R., Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigglio,J., Vassiliev,H., Viel,R., Vo.A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zimmer,A. and Zody,M.

TITLE Direct Submission
JOURNAL Submitted (14-FEB-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT On Jul 4, 2001 this sequence version replaced gi:7139677.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence.submissions@genome.wi.mit.edu

----- Project Information

Center project name: L4260

Center clone name: 687_M23

----- Summary Statistics

Sequencing vector: M13; M77815; 47% of reads

Sequencing vector: plasmid; n/a; 53% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 188836 bases at least Q40

Consensus quality: 189476 bases at least Q30

Consensus quality: 189595 bases at least Q20

Insert size: 188000; agarose-fp

Insert size: 189663; sum-of-contigs

Quality coverage: 10.7 in Q20 bases; agarose-fp

Quality coverage: 10.7 in Q20.

* NOTE: This is a 'working draft' sequence. It currently

* consists of 4 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

* as soon as it is available and the accession number will

* be preserved.

* 1 37424: contig of 37424 bp in length

* 37425 37524: gap of 100 bp

* 37525 42149: contig of 4625 bp in length

* 42150 42249: gap of 100 bp

* 42250 177306: contig of 135057 bp in length

* 177307 177406: gap of 100 bp

* 177407 189963: contig of 12557 bp in length.

FEATURES

source

1. .189963

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="8"

/map="8"

/clone="RP11-687M23"

/clone_lib="RPCI-11 Human Male BAC"

1. .37424

/notes="assembly_fragment"

clone_end:SP6

vector_side:left"

37425. .37524

/estimated_length=100

37525. .42149

misc_feature

gap

misc_feature

gap		/estimated_length=unknown		COMMENT	
12673..12772		12673..12772		On Aug 21, 2002 this sequence version replaced gi:22002697.	
gap		/estimated_length=unknown		----- Genome Center	
13890..13989		13890..13989		Center: Wellcome Trust Sanger Institute	
gap		/estimated_length=unknown		Web site: http://www.sanger.ac.uk	
16329..16428		16329..16428		Contact: humquery@sanger.ac.uk	
gap		/estimated_length=unknown		-----	
18332..18431		18332..18431		During sequence assembly data is compared from overlapping clones.	
gap		/estimated_length=unknown		Where differences are found these are annotated as variations	
20563..20662		20563..20662		together with a note of the overlapping clone name. Note that the	
gap		/estimated_length=unknown		variation annotation may not be found in the sequence submission	
22800..22899		22800..22899		corresponding to the overlapping clone, as we submit sequences with	
gap		/estimated_length=unknown		only a small overlap as described above.	
24689..24788		24689..24788		This sequence was finished as follows unless otherwise noted: all	
gap		/estimated_length=unknown		regions were either double-stranded or sequenced with an alternate	
28948..29047		28948..29047		chemistry or covered by high quality data (i.e., phred quality >=	
gap		/estimated_length=unknown		30); an attempt was made to resolve all sequencing problems, such	
33637..33736		33637..33736		as compressions and repeats; all regions were covered by at least	
gap		/estimated_length=unknown		one plasmid subclone or more than one M13 subclone; and the	
37009..37108		37009..37108		assembly was confirmed by restriction digest. The following	
gap		/estimated_length=unknown		abbreviations are used to associate primary accession numbers given	
42486..42585		42486..42585		in the feature table with their source databases: Em: EMBL; Sw:	
gap		/estimated_length=unknown		SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP	
50240..50339		50240..50339		database can be found at	
gap		/estimated_length=unknown		http://www.sanger.ac.uk/projects/C.elegans/wormpep RP23-229B23 is	
58128..58227		58128..58227		from the RPCI-23 Mouse PAC Library	
gap		/estimated_length=unknown		constructed by the group of Pieter de Jong.	
65777..65876		65777..65876		For further details see http://www.chori.org/bacpac/home.htm	
75168..75267		75168..75267		VECTOR: pBACes.6. Location/Qualifiers	
gap		/estimated_length=unknown		1. .210718	
84182..84281		84182..84281		/organism="Mus musculus"	
gap		/estimated_length=unknown		/mol_type="genomic DNA"	
97294..97393		97294..97393		/db_xref="taxon:10090"	
116123..116222		116123..116222		/chromosome="4"	
gap		/estimated_length=unknown		/clone="RP23-229B23"	
136123..136222		136123..136222		/clone_lib="RPCI-23"	
gap		/estimated_length=unknown		ORIGIN	
159177..159276		159177..159276		Query Match 80.8%; Score 20.2; DB 9; Length 210718;	
gap		/estimated_length=unknown		Best Local Similarity 88.0%; Pred. No. 1.2e+02;	
183074..183173		183074..183173		Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;	
/estimated_length=unknown		/estimated_length=unknown		QY 1 AAAAAACAATGACCACTGGGCTGT 25	
ORIGIN		ORIGIN		DB 54146 AAGAACAAAGACCACTGAGCTGT 54170	
Query Match 80.8%; Score 20.2; DB 14; Length 208161;		Query Match 80.8%; Score 20.2; DB 14; Length 208161;		RESULT 26	
Best Local Similarity 88.0%; Pred. No. 1.2e+02;		Best Local Similarity 88.0%; Pred. No. 1.2e+02;		AC094552/c	
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;		Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;		AC094552/LOCUS	
QY 1 AAAAAACAATGACCACTGGGCTGT 25		QY 1 AAAAAACAATGACCACTGGGCTGT 25		DEFINITION	
DB 189749 AAAAAAGATGAGAACTGGGCTGT 189725		DB 189749 AAAAAAGATGAGAACTGGGCTGT 189725		AC094552.7 GI:30466691	
RESULT 25		RESULT 25		VERSION	
AL683818		AL683818		KEYWORDS	
LOCUS		LOCUS		SOURCE	
DEFINITION		DEFINITION		ORGANISM	
Mouse DNA sequence from clone RP23-229B23 on chromosome 4, complete		Mouse DNA sequence from clone RP23-229B23 on chromosome 4, complete		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
sequence.		sequence.		Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;	
AL683818		AL683818		Sciurognathi; Muroidae; Muridae; Murinae; Rattus.	
HTG.		HTG.		1 (bases 1 to 210718)	
Mus musculus (house mouse)		Mus musculus (house mouse)		REFERENCE	
Mus musculus		Mus musculus		AUTHORS	
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		TITLE	
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;		Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;		JOURNAL	
Sciurognathi; Muroidae; Muridae; Murinae; Mus.		Sciurognathi; Muroidae; Muridae; Murinae; Mus.		humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk	
1 (bases 1 to 210718)		1 (bases 1 to 210718)			
Dunn.M.		Dunn.M.			
Direct Submission		Direct Submission			
Submitted (24-AUG-2002) Wellcome Trust Sanger Institute, Hinxton,		Submitted (24-AUG-2002) Wellcome Trust Sanger Institute, Hinxton,			
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:		Cambridgeshire, CB10 1SA, UK. E-mail enquiries:			
humquery@sanger.ac.uk		humquery@sanger.ac.uk			

Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Frazer, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregiorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Haylak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowale, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensuewa, L., Louseged, H., Lozado, R.J., Lu, X., Ma, J., Mangum, B., Mapua, P., Martin, K., Mahmoud, M., Malloy, K., Mangum, A., Mahoney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Parks, K., Nwokedemeh, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajls, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steinkle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villaseana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, Q., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczek, R., Woodson, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, P., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausen, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

Direct Submission
Unpublished
2 (bases 1 to 246315)
Worley, K.C.
Direct Submission
Submitted (15-SEP-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 246315)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (09-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On May 9, 2003 this sequence version replaced gi:24818682.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

Center: Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GAWW
Center clone name: CH230-4p9

----- Summary Statistics
Assembly program: Atlas
Consensus quality: 197283 bases at least Q40
Consensus quality: 20460 bases at least Q30
Consensus quality: 206455 bases at least Q20
Estimated insert size: 211719; sum-of-contigs estimation
Quality coverage: 5x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
* as soon as it is available and the accession number will
* be preserved.
* 1 242724: contig of 242724 bp in length
* 242725 242824: gap of unknown length
* 242825 243872: contig of 1048 bp in length
* 243873 243972: gap of unknown length
* 243973 245076: contig of 1104 bp in length
* 245077 245176: gap of unknown length
* 245177 246315: contig of 1139 bp in length.

FEATURES
source
1. 246315
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-4p9"
1. 1812
/note="wgs_end_extension"
Clone end: T7
3382. 4325
/note="clone_boundary"
clone end: T7
site: EcoRI
end_sequence: BH310992"
25124. 27395
/note="wgs_contig"
107600. 108656
/note="wgs_contig"
complement(241125..241757)
/note="clone_boundary"
clone_end: Sp6
site: EcoRI
end_sequence: BH310993"
242725. 242824
/estimated_length=unknown
243873. 243972
/estimated_length=unknown
245077. 245176
/estimated_length=unknown
ORIGIN
Query Match 80.8%; Score 20.2; DB 14; Length 246315;
Best Local Similarity 88.0%; Pred. No. 1.2e+02;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 AAAAAACAATGACCAACTGGGCTGT 25
||||||| ||||| ||||| ||||| |||||
Db 56873 AAAAAACAATGACCAACTGGGCTGT 56849
RESULT 27
AC094299
LOCUS
DEFINITION Rattus norvegicus clone CH230-314, *** SEQUENCING IN PROGRESS ***,
4 unordered pieces.
AC094299
AC094299.9 GI:30467425
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.

SOURCE	Rattus norvegicus (Norway rat)
ORGANISM	Rattus norvegicus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Muridae; Murinae; Rattus.
AUTHORS	1 (bases 1 to 249661) Muzny, D. Marie., Metzker, M. Lee., Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denison, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Evans, K., Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hoques, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, B., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpach, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, D., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensuhewa, L., Louisedge, H., Lozada, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhinney, S., McLeod, M. P., McNeill, T. Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwaokemele, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkoch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L., L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smajls, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wlezyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstock, G. and Gibbs, R. A.
REFERENCE	2 (bases 1 to 249661) Worley, K. C.
TITLE	Direct Submission
JOURNAL	Submitted (16-SEP-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE	3 (bases 1 to 249661) Rat Genome Sequencing Consortium.
TITLE	Direct Submission
JOURNAL	Submitted (09-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
COMMENT	On May 9, 2003 this sequence version replaced gi:23322519. The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described

in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GAIT
Center clone name: CH230-314
----- Summary Statistics
Assembly program: Atlas
Consensus quality: 219534 bases at least Q40
Consensus quality: 221593 bases at least Q30
Consensus quality: 222699 bases at least Q20
Estimated insert size: 234897; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 243876: contig of 243876 bp in length
* 243877 243976: gap of unknown length
* 243977 245492: contig of 1516 bp in length
* 245493 245592: gap of unknown length
* 245593 246913: contig of 1321 bp in length
* 246914 247013: gap of unknown length
* 247014 249661: contig of 2648 bp in length.
FEATURES
source
1. .249661
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-314"
1. .2422
/notes="wgs end _extension
clone end:T7"
complement(13592..14341)
/note="clone_boundary
clone end:T7
site:ECORI
end sequence:BH311293"
24207..224703
/notes="clone_boundary
clone end:Sp6
site:ECORI
end sequence:BH311296"
225324..226449
/notes="wgs end _extension
clone end:Sp6"
232292..233410
/notes="wgs end _extension
clone end:Sp6"
241777..243876
/notes="wgs end _extension
clone end:Sp6"
243877..243976
/estimated_length=unknown
245493..245592
misc_feature
misc_feature
misc_feature
misc_feature
misc_feature
misc_feature
gap
gap

gap /estimated_length=unknown
246914..247013
/estimated_length=unknown

ORIGIN

Query Match 80.8%; Score 20.2; DB 14; Length 249661;
Best Local Similarity 88.0%; Pred. No. 1.2e+02;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAACAATGACCAACTGGGTGT 25
||||||| ||||| ||||| ||||| |||||

Db 15854 AAAAAACAATGACCAACTGAGCTGT 15878
||||||| ||||| ||||| ||||| |||||

RESULT 28
AC146241 160620 bp DNA linear PRI 30-JUN-2004
DEFINITION Pan troglodytes BAC clone RP43-31L17 from 7, complete sequence.
AC146241
AC146241.4 GI:48958798
VERSION HTG.
KEYWORDS Pan troglodytes (chimpanzee)
SOURCE Pan troglodytes
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Pan.
REFERENCE 1 (bases 1 to 160620)
AUTHORS Tomlinson, C. and Haglund, K.
TITLE The sequence of Pan troglodytes BAC clone RP43-31L17
JOURNAL Unpublished (2001)
REFERENCE 2 (bases 1 to 160620)
AUTHORS Wilson, R.K.
TITLE Direct Submission
JOURNAL Submitted (01-AUG-2003) Genetics, Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
REFERENCE 3 (bases 1 to 160620)
AUTHORS Wilson, R.K.
TITLE Direct Submission
JOURNAL Submitted (12-MAY-2004) Genetics, Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
REFERENCE 4 (bases 1 to 160620)
AUTHORS Wilson, R.K.
TITLE Direct Submission
JOURNAL Submitted (19-JUN-2004) Genetics, Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
REFERENCE 5 (bases 1 to 160620)
AUTHORS Wilson, R.K.
TITLE Direct Submission
JOURNAL Submitted (30-JUN-2004) Washington University School of Medicine, Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
COMMENT On Jun 19, 2004 this sequence version replaced gi:47131419.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu
Contact: submissions@genome.wustl.edu
----- Summary Statistics
Center project name: C_PT031L17

NOTICE:

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. Wes Warren,

Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu>

SOURCE INFORMATION:

The RPCI-43 BAC Library has been constructed by Chung-Li Shu. DNA was isolated from white blood cells obtained from a male chimpanzee (Pan troglodytes, 'Clint', Yerkes #C0471; birthdate: 6-6-80). The clone and detailed information can be obtained from ResGen (<http://www.resgen.com>) or Pieter de Jong and co-workers at <http://www.bacpac.chori.org>.

NEIGHBORING SEQUENCE INFORMATION:

This sequence is the entire insert of the clone. This clone is overlapped by AC140950.

FEATURES
source

Location/Qualifiers
1..160620
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/chromosome="7"
/map="7"
/clone="RP43-31L17"
/clone_lib="RPCI-43"

ORIGIN

Query Match 79.2%; Score 19.8; DB 8; Length 160620;
Best Local Similarity 91.3%; Pred. No. 1.9e+02;
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AAAAAACAATGACCAACTGGGTGT 24
||||||| ||||| ||||| ||||| |||||

Db 63994 AAAAAACAATGACCAACAGGATG 64016

RESULT 29

AC166995/c

LOCUS

DEFINITION

AC166995

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Center: Washington University Genome Sequencing Center

Center code: WUGSC

Web site: http://genome.wustl.edu

Contact: submissions@genome.wustl.edu

----- Summary Statistics

Center project name: M_BA0255N15

Sequencing vector: pM13; 0%

Sequencing method: 100%

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.990319

Consensus quality: 161427 bases at least Q40

Consensus quality: 163374 bases at least Q30

Sequencing vector: pM13; 0%

Sequencing method: 100%

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.990319

Consensus quality: 161427 bases at least Q40

Consensus quality: 163374 bases at least Q30

Sequencing vector: pM13; 0%

Sequencing method: 100%

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.990319

Consensus quality: 161427 bases at least Q40

Consensus quality: 163374 bases at least Q30

Consensus quality: 164740 bases at least Q20

* NOTE: This is a 'working draft' sequence. It currently
* consists of 17 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1251: contig of 1251 bp in length
* 1252 1351: gap of unknown length
* 1352 3283: contig of 1932 bp in length
* 3284 3383: gap of unknown length
* 3384 4641: contig of 1258 bp in length
* 4642 4741: gap of unknown length
* 4742 8114: contig of 3373 bp in length
* 8115 8214: gap of unknown length
* 8215 13181: contig of 4967 bp in length
* 13182 13281: gap of unknown length
* 13282 16008: contig of 2727 bp in length
* 16009 16108: gap of unknown length
* 16109 21115: contig of 5007 bp in length
* 21116 21215: gap of unknown length
* 21216 24744: contig of 3529 bp in length
* 24745 24844: gap of unknown length
* 24845 32255: contig of 7411 bp in length
* 32256 32356: gap of unknown length
* 32357 40227: contig of 7871 bp in length
* 40228 40327: gap of unknown length
* 40328 47032: contig of 6706 bp in length
* 47033 47133: gap of unknown length
* 47134 55420: contig of 8288 bp in length
* 55421 55521: gap of unknown length
* 55522 66858: contig of 11338 bp in length
* 66859 79315: gap of unknown length
* 79316 79416: gap of unknown length
* 79417 98104: contig of 18688 bp in length
* 98105 98204: gap of unknown length
* 98205 133431: contig of 35227 bp in length
* 133432 133530: gap of unknown length
* 133531 170524: contig of 36994 bp in length.

FEATURES

source
1. .170524
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="16"
/clone="RP23-255N15"
misc_feature
1. .1251
/note="assembly_name:Contig14"
gap
1252. .1351
/estimated_length=unknown
misc_feature
1352. .3283
/note="assembly_name:Contig15"
gap
3284. .3383
/estimated_length=unknown
misc_feature
3384. .4641
/note="assembly_name:Contig16"
gap
4642. .4741
/estimated_length=unknown
misc_feature
4742. .8114
/note="assembly_name:Contig18"
gap
8115. .8214
/estimated_length=unknown
misc_feature
8215. .13181
/note="assembly_name:Contig19"
gap
13182. .13281
/estimated_length=unknown
misc_feature
13282. .16008
/note="assembly_name:Contig20"
gap
16009. .16108

misc_feature
16109. .21115
/note="assembly_name:Contig21"
gap
21116. .21215
/estimated_length=unknown
misc_feature
21216. .24744
/note="assembly_name:Contig22"
gap
24745. .24844
/estimated_length=unknown
misc_feature
24845. .32255
/note="assembly_name:Contig24"
gap
32256. .32355
/estimated_length=unknown
misc_feature
32356. .40226
/note="assembly_name:Contig25"
gap
40227. .40326
/estimated_length=unknown
misc_feature
40327. .47032
/note="assembly_name:Contig26"
gap
47033. .47132
/estimated_length=unknown
misc_feature
47133. .55420
/note="assembly_name:Contig27"
gap
55421. .55520
/estimated_length=unknown
misc_feature
55521. .66858
/note="assembly_name:Contig28"
gap
66859. .66958
/estimated_length=unknown
misc_feature
66959. .79315
/note="assembly_name:Contig29
clone_end:T7
vector_side:left"
gap
79316. .79415
/estimated_length=unknown
misc_feature
79416. .98103
/note="assembly_name:Contig30"
gap
98104. .98203
/estimated_length=unknown
misc_feature
98204. .133430
/note="assembly_name:Contig31
clone_end:SP6
vector_side:right"
gap
133431. .133530
/estimated_length=unknown
misc_feature
133531. .170524
/note="assembly_name:Contig32"

ORIGIN

Query Match 79.2%; Score 19.8; DB 14; Length 170524;
Best Local Similarity 91.3%; Pred. No. 1.9e+02;
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 AAAAAACAATGACCAACTGGGCT 23
|||||
Db 161835 AAAAAACAATGACCACTGGCT 161813

RESULT 30

AC073470
LOCUS
DEFINITION Homo sapiens chromosome 7 clone RP11-570M16, WORKING DRAFT
AC073470 171673 bp DNA linear HTG 04-AUG-2000
SEQUENCE, 33 unordered pieces.
ACCESSION AC073470
VERSION AC073470.3 GI:9690412
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 171673)
AUTHORS Waterston,R.H.

TITLE The sequence of Homo sapiens clone
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 171673)
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (18-JUN-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
COMMENT On Aug 4, 2000 this sequence version replaced gi:9309529.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu/gsc/index.shtml>
----- Project Information -----
Center project name: H.NH0570M16
----- Summary Statistics -----
Sequencing vector: M13; 100%
Sequencing vector: plasmid; 0%
Chemistry: Dye-terminator Big Dye; 0% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 151549 bases at least Q40
Consensus quality: 157491 bases at least Q30
Consensus quality: 160825 bases at least Q20
Insert size: 168000; agarose-fp
Insert size: 168473; sum-of-contigs
Quality coverage: 3.65 in Q20 bases; agarose-fp
Quality coverage: 3.73 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 33 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1
* 1043: contig of 1043 bp in length
* 1044: 1143: gap of unknown length
* 1144: 2412: contig of 1269 bp in length
* 2413: 2512: gap of unknown length
* 2513: 3933: contig of 1421 bp in length
* 3934: 4033: gap of unknown length
* 4034: 5390: contig of 1357 bp in length
* 5391: 5490: gap of unknown length
* 5491: 6629: contig of 1139 bp in length
* 6630: 6729: gap of unknown length
* 6730: 7966: contig of 1237 bp in length
* 7967: 8066: gap of unknown length
* 8067: 9485: contig of 1419 bp in length
* 9486: 9585: gap of unknown length
* 9586: 10897: contig of 1312 bp in length
* 10898: 12498: contig of 1501 bp in length
* 12499: 12598: gap of unknown length
* 12599: 13735: contig of 1137 bp in length
* 13736: 13835: gap of unknown length
* 13836: 15277: contig of 1442 bp in length
* 15278: 15377: gap of unknown length
* 15378: 17363: contig of 1986 bp in length
* 17364: 17463: gap of unknown length
* 17464: 18722: contig of 1259 bp in length
* 18723: 18822: gap of unknown length
* 18823: 21225: contig of 2403 bp in length
* 21226: 21325: gap of unknown length
* 21326: 23191: contig of 1866 bp in length
* 23192: 23291: gap of unknown length
* 23292: 26263: contig of 2972 bp in length
* 26264: 26363: gap of unknown length
* 26364: 30191: contig of 3828 bp in length
* 30192: 30291: gap of unknown length
* 30292: 32484: contig of 2193 bp in length

* 32485 32584: gap of unknown length
* 32585 36044: contig of 3460 bp in length
* 36045 36144: gap of unknown length
* 36145 41434: contig of 5290 bp in length
* 41435 41534: gap of unknown length
* 41535 41533: contig of 2999 bp in length
* 44534 44633: gap of unknown length
* 44634 50067: contig of 5434 bp in length
* 50068 50167: gap of unknown length
* 50168 55034: contig of 4867 bp in length
* 55035 55134: gap of unknown length
* 55135 60252: contig of 5118 bp in length
* 60253 60352: gap of unknown length
* 60353 67994: contig of 7642 bp in length
* 67995 68094: gap of unknown length
* 68095 74834: contig of 6740 bp in length
* 74835 74934: gap of unknown length
* 74935 83346: contig of 8411 bp in length
* 83346 83446: gap of unknown length
* 83446 95023: contig of 11578 bp in length
* 95024 95123: gap of unknown length
* 95124 103633: contig of 8510 bp in length
* 103634 103733: gap of unknown length
* 103734 115298: contig of 11565 bp in length
* 115299 115398: gap of unknown length
* 115399 128887: contig of 13489 bp in length
* 128888 128987: gap of unknown length
* 128988 145911: contig of 18924 bp in length
* 145912 146011: gap of unknown length
* 146012 171673: contig of 25662 bp in length.

FEATURES

Location/Qualifiers
1..171673
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="7"
/clone="RP11-570M16"
1..1043
/note="assembly_name:Contig14"
1044..1143
/estimated_length=unknown
1144..2412
/note="assembly_name:Contig28"
2413..2512
/estimated_length=unknown
2513..3933
/note="assembly_name:Contig35"
3934..4033
/estimated_length=unknown
4034..5390
/note="assembly_name:Contig36"
5391..5490
/estimated_length=unknown
5491..6629
/note="assembly_name:Contig37"
6630..6729
/estimated_length=unknown
6730..7966
/note="assembly_name:Contig38"
7967..8066
/estimated_length=unknown
8067..9485
/note="assembly_name:Contig39"
9486..9585
/estimated_length=unknown
9586..10897
/note="assembly_name:Contig40"
10898..10997
/estimated_length=unknown
10998..12498
/note="assembly_name:Contig41"
12499..12598
/estimated_length=unknown

5839	5638	481	<800	915	928	2590	2612	23	<800	7123	7037
15389	15229	2871	2817	2190	2318	6365	6393	3713	3799	2315	2229
606	<800	1314	1322	133	<800	714	<800	676	<800	6908	7037
3819	3880	25	<800	223	<800	5839	5956	5582	5360	90	<800
13873	13714	7045	7062	1621	1636	669	<800	796	<800	12239	12115
1176	1163	2024	2044	8577	8641	317	<800	1417	1322	1168	1185
945	959	2748	2817	6552	6504	2638	2612	1330	1322	5227	5246
2616	2612	40	<800	3953	3971	8082	8138	3827	3799	2879	2850
2208	2285	1427	1395	6996	7037	493	<800	7059	7062	689	<800
1422	1381	4849	4848	132	<800	6654	6682	11906	11857	4381	4512
1734	1728	2499	2492	2240	2229	78	<800	5834	5772	2454	2463
1773	1728	7466	7511	4528	4512	2276	2410	2597	2607	266	<800
758	<800	8699	8670	2449	2463	3810	3708	600	<800	4188	4163
4545	4468	2816	2817	6606	6504	81	<800	499	<800	3383	3415
1858	1866	1637	1544	3886	3807	8	<800	2176	2214	4129	4163
2091	2208	585	<800	1378	1356	---	---	85	<800	175	<800
7187	7164	107	<800	3587	3634	---	---	2192	2214	---	---
2138	2285	3758	3799	3099	3103	Query Match 79.2%; Score 19.8; DB 8; Length 181935;					
315	<800	1505	1544	4601	4512	Best Local Similarity 91.3%; Pred. No. 1.9e+02;					
12337	12706	2265	2350	132	<800	Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;					
1362	1381	2796	2817	2636	2652	QY 2 AAAAAAATGACCAACTGGCTG 24					
1222	1228	1173	1139	1669	1636						
8377	8459	2310	2350	4012	4163	Db 175449 AAAAAAATGACCAACAGGATG 175427					
3664	3708	512	<800	6308	6302	AC140950 190084 bp DNA linear PRI 26-SEP-2003					
5399	5445	9486	9423	2024	2047	Pan troglodytes BAC clone RP43-58122 from 7, complete sequence.					
4687	4639	2326	2350	2608	2579	AC140950.1 GI:28850252					
2045	2099	3865	3799	2455	2463	HTG.					
972	959	3502	3511	47	<800	Pan troglodytes (chimpanzee)					
295	<800	5390	5567	13348	13324	Pan troglodytes					
2427	2612	2803	2817	2744	2730	Pan troglodytes					
1085	1078	900	904	638	<800	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
2591	2612	881	904	1520	1509	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;					
2275	2285	6801	6738	3270	3220	Hominidae; Pan.					
2414	2410	5345	5360	7440	7465	1 (bases 1 to 190084)					
3719	3708	1833	1756	5039	4961	Walligorski, J., Haakenson, W. and Haglund, K.					
5636	5445	1059	1060	4761	4747	The sequence of Pan troglodytes BAC clone RP43-58122					
						Unpublished (2001)					
						2 (bases 1 to 190084)					
						Sulston, J.E. and Wilson, R.					
						Sequencing of Pan troglodytes					
						Unpublished (2001)					
						3 (bases 1 to 190084)					
						Waterston, R.					
						Direct Submission					
						Submitted (05-MAR-2003) Department of Genetics, Washington					
						University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA					
						4 (bases 1 to 190084)					
						Wilson, R.					
						Direct Submission					
						Submitted (26-SEP-2003) Department of Genetics, Washington					

COMMENT

University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu>
Contact: submissions@wustl.edu
----- Summary Statistics
Center project name: C_PT058122

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. Wes Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu>

SOURCE INFORMATION:
The RPCI-43 BAC Library has been constructed by Chung-Li Shu. DNA was isolated from white blood cells obtained from a male chimpanzee (Pan troglodytes, 'Clint', Yerkes #C0471; birthdate: 6-6-80). The clone and detailed information can be obtained from ResGen (<http://www.resgen.com>) or Pieter de Jong and co-workers at <http://www.bacpac.chori.org>.

NEIGHBORING SEQUENCE INFORMATION:
This sequence is the entire insert of the clone.

FEATURES
source : 1.190084
 /organism="Pan troglodytes"
 /mol_type="genomic DNA"
 /db_xref="taxon:9598"
 /chromosome="7"
 /map="7"
 /clone="RP43-58122"
 /clone_lib="RPCI-43"
 106485..106511
misc_feature /note="Unresolved homopolymeric repeat."

ORIGIN

Query Match 79.2%; Score 19.8; DB 8; Length 190084;
Best Local Similarity 91.3%; Pred. No. 1.9e+02;
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 AAAACAATGACCACTGGGCTG 24
|||||
Db 55534 AAAACAATGACCAACAGGATG 55556

RESULT 33

CT010460
LOCUS CT010460 232469 bp DNA linear HTG 25-JUL-2005
DEFINITION Mus musculus chromosome 12 clone RP23-311G21, *** SEQUENCING IN PROGRESS ***, 16 unordered pieces.
ACCESSION CT010460
VERSION CT010460.1 GI:71142028
KEYWORDS HTG; HTGS PHASE1.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Sciurognathi; Muroidae; Muridae; Murinae; Mus.
1 (bases 1 to 232469)
Sims, S.
Direct Submission
Submitted (21-JUL-2005) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vega@sanger.ac.uk
Clone requests: clonerequest@sanger.ac.uk
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: <http://www.sanger.ac.uk>
Contact: vega@sanger.ac.uk
----- Project Information
Center project name: BM311G21
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; 100% of reads
Consensus quality: 226307 bases at least Q40
Consensus quality: 227941 bases at least Q30
Consensus quality: 229053 bases at least Q20
Insert size: 230969; sum-of-contigs
Insert size: 243194; 4.0% error; agarose-fp
Quality coverage: 4.52x in Q20 bases; sum-of-contigs Quality coverage: 4.41x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently consists of 16 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. * This record will be updated with the finished sequence * as soon as it is available and the accession number will be preserved.

1 2300: contig of 2300 bp in length
2301 2400: gap of 100 bp
2401 20064: contig of 17664 bp in length
20065 20164: gap of 100 bp
20165 32528: contig of 12364 bp in length
32529 32529: gap of 100 bp
32530 50388: contig of 17760 bp in length
50389 50389: gap of 100 bp
50390 56605: contig of 6117 bp in length
56606 56705: gap of 100 bp
56706 72035: contig of 15330 bp in length
72036 72135: gap of 100 bp
72136 114525: contig of 42390 bp in length
114526 114625: gap of 100 bp
114626 126383: contig of 11758 bp in length
126384 126483: gap of 100 bp
126484 142084: contig of 15601 bp in length
142085 142184: gap of 100 bp
142185 152571: contig of 10387 bp in length
152572 152671: gap of 100 bp
152672 155785: contig of 3114 bp in length
155786 155885: gap of 100 bp
155886 170998: contig of 15113 bp in length
170999 171098: gap of 100 bp
171099 175427: contig of 4329 bp in length
175428 175527: gap of 100 bp
175528 203780: contig of 28253 bp in length
203781 203880: gap of 100 bp
203881 206442: contig of 2562 bp in length
206443 206542: gap of 100 bp
206543 232469: contig of 25927 bp in length.

FEATURES
source Location/Qualifiers
1..232469
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /chromosome="12"
 /clone="RP23-311G21"
 /clone_lib="RPCI-23"
misc_feature 1..2300

Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F.,
Takaku-Akaiira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K.,
Yasunishi, A. and Hayashizaki, Y.

FEATURES

Location/Qualifiers
1..2346
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="J023133N22"

ORIGIN

Query Match 77.6%; Score 19.4; DB 15; Length 2346;
Best Local Similarity 95.4%; Pred. No. 3.2e+02;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 AAAACAATGACCAACTGGGCT 23

Db 1483 AAATAATGACCAACTGGGCT 1463

RESULT 35

AP008212_261/c

WPCOMMENT

Sequence split into 308 fragments LOCUS AP008212 Accession AP008212

Fragment Name	Begin	End
AP008212_000	1	110000
AP008212_001	100001	210000
AP008212_002	200001	310000
AP008212_003	300001	410000
AP008212_004	400001	510000
AP008212_005	500001	610000
AP008212_006	600001	710000
AP008212_007	700001	810000
AP008212_008	800001	910000
AP008212_009	900001	1010000
AP008212_010	1000001	1110000
AP008212_011	1100001	1210000
AP008212_012	1200001	1310000
AP008212_013	1300001	1410000
AP008212_014	1400001	1510000
AP008212_015	1500001	1610000
AP008212_016	1600001	1710000
AP008212_017	1700001	1810000
AP008212_018	1800001	1910000
AP008212_019	1900001	2010000
AP008212_020	2000001	2110000
AP008212_021	2100001	2210000
AP008212_022	2200001	2310000
AP008212_023	2300001	2410000
AP008212_024	2400001	2510000
AP008212_025	2500001	2610000
AP008212_026	2600001	2710000
AP008212_027	2700001	2810000
AP008212_028	2800001	2910000
AP008212_029	2900001	3010000
AP008212_030	3000001	3110000
AP008212_031	3100001	3210000
AP008212_032	3200001	3310000
AP008212_033	3300001	3410000
AP008212_034	3400001	3510000
AP008212_035	3500001	3610000
AP008212_036	3600001	3710000
AP008212_037	3700001	3810000
AP008212_038	3800001	3910000
AP008212_039	3900001	4010000
AP008212_040	4000001	4110000
AP008212_041	4100001	4210000
AP008212_042	4200001	4310000
AP008212_043	4300001	4410000
AP008212_044	4400001	4510000
AP008212_045	4500001	4610000

AP008212_046	4600001	4710000
AP008212_047	4700001	4810000
AP008212_048	4800001	4910000
AP008212_049	4900001	5010000
AP008212_050	500001	5110000
AP008212_051	5100001	5210000
AP008212_052	5200001	5310000
AP008212_053	5300001	5410000
AP008212_054	5400001	5510000
AP008212_055	5500001	5610000
AP008212_056	5600001	5710000
AP008212_057	5700001	5810000
AP008212_058	5800001	5910000
AP008212_059	5900001	6010000
AP008212_060	6000001	6110000
AP008212_061	6100001	6210000
AP008212_062	6200001	6310000
AP008212_063	6300001	6410000
AP008212_064	6400001	6510000
AP008212_065	6500001	6610000
AP008212_066	6600001	6710000
AP008212_067	6700001	6810000
AP008212_068	6800001	6910000
AP008212_069	6900001	7010000
AP008212_070	7000001	7110000
AP008212_071	7100001	7210000
AP008212_072	7200001	7310000
AP008212_073	7300001	7410000
AP008212_074	7400001	7510000
AP008212_075	7500001	7610000
AP008212_076	7600001	7710000
AP008212_077	7700001	7810000
AP008212_078	7800001	7910000
AP008212_079	7900001	8010000
AP008212_080	8000001	8110000
AP008212_081	8100001	8210000
AP008212_082	8200001	8310000
AP008212_083	8300001	8410000
AP008212_084	8400001	8510000
AP008212_085	8500001	8610000
AP008212_086	8600001	8710000
AP008212_087	8700001	8810000
AP008212_088	8800001	8910000
AP008212_089	8900001	9010000
AP008212_090	9000001	9110000
AP008212_091	9100001	9210000
AP008212_092	9200001	9310000
AP008212_093	9300001	9410000
AP008212_094	9400001	9510000
AP008212_095	9500001	9610000
AP008212_096	9600001	9710000
AP008212_097	9700001	9810000
AP008212_098	9800001	9910000
AP008212_099	9900001	10010000
AP008212_100	10000001	10110000
AP008212_101	10100001	10210000
AP008212_102	10200001	10310000
AP008212_103	10300001	10410000
AP008212_104	10400001	10510000
AP008212_105	10500001	10610000
AP008212_106	10600001	10710000
AP008212_107	10700001	10810000
AP008212_108	10800001	10910000
AP008212_109	10900001	11010000
AP008212_110	11000001	11110000
AP008212_111	11100001	11210000
AP008212_112	11200001	11310000
AP008212_113	11300001	11410000
AP008212_114	11400001	11510000
AP008212_115	11500001	11610000
AP008212_116	11600001	11710000
AP008212_117	11700001	11810000
AP008212_118	11800001	11910000

AP008212_119 11900001 12010000
AP008212_120 12000001 12110000
AP008212_121 12100001 12210000
AP008212_122 12200001 12310000
AP008212_123 12300001 12410000
AP008212_124 12400001 12510000
AP008212_125 12500001 12610000
AP008212_126 12600001 12710000
AP008212_127 12700001 12810000
AP008212_128 12800001 12910000
AP008212_129 12900001 13010000
AP008212_130 13000001 13110000
AP008212_131 13100001 13210000
AP008212_132 13200001 13310000
AP008212_133 13300001 13410000
AP008212_134 13400001 13510000
AP008212_135 13500001 13610000
AP008212_136 13600001 13710000
AP008212_137 13700001 13810000
AP008212_138 13800001 13910000
AP008212_139 13900001 14010000
AP008212_140 14000001 14110000
AP008212_141 14100001 14210000
AP008212_142 14200001 14310000
AP008212_143 14300001 14410000
AP008212_144 14400001 14510000
AP008212_145 14500001 14610000
AP008212_146 14600001 14710000
AP008212_147 14700001 14810000
AP008212_148 14800001 14910000
AP008212_149 14900001 15010000
AP008212_150 15000001 15110000
AP008212_151 15100001 15210000
AP008212_152 15200001 15310000
AP008212_153 15300001 15410000
AP008212_154 15400001 15510000
AP008212_155 15500001 15610000
AP008212_156 15600001 15710000
AP008212_157 15700001 15810000
AP008212_158 15800001 15910000
AP008212_159 15900001 16010000
AP008212_160 16000001 16110000
AP008212_161 16100001 16210000
AP008212_162 16200001 16310000
AP008212_163 16300001 16410000
AP008212_164 16400001 16510000
AP008212_165 16500001 16610000
AP008212_166 16600001 16710000
AP008212_167 16700001 16810000
AP008212_168 16800001 16910000
AP008212_169 16900001 17010000
AP008212_170 17000001 17110000
AP008212_171 17100001 17210000
AP008212_172 17200001 17310000
AP008212_173 17300001 17410000
AP008212_174 17400001 17510000
AP008212_175 17500001 17610000
AP008212_176 17600001 17710000
AP008212_177 17700001 17810000
AP008212_178 17800001 17910000
AP008212_179 17900001 18010000
AP008212_180 18000001 18110000
AP008212_181 18100001 18210000
AP008212_182 18200001 18310000
AP008212_183 18300001 18410000
AP008212_184 18400001 18510000
AP008212_185 18500001 18610000
AP008212_186 18600001 18710000
AP008212_187 18700001 18810000
AP008212_188 18800001 18910000
AP008212_189 18900001 19010000
AP008212_190 19000001 19110000
AP008212_191 19100001 19210000

AP008212_192 19200001 19310000
AP008212_193 19300001 19410000
AP008212_194 19400001 19510000
AP008212_195 19500001 19610000
AP008212_196 19600001 19710000

Query Match 77.6%; Score 19.4; DB 15; Length 110000;
Best Local Similarity 95.2%; Pred. No. 2.9e+02;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 AAACAATGACCACTGGGCT 23
|||||
Db 52673 AAAATAATGACCACTGGGCT 52653

RESULT 36

AP003517/c

LOCUS

DEFINITION

AP003517

VERSION

AP003517.3

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AP003517 151636 bp DNA linear PIN 25-AUG-2004
Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 6,
BAC clone:OSUNBa0062J02.

AP003517
AP003517.3 GI:51535415

Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.

1 Sasaki, T., Matsumoto, T. and Yamamoto, K.
Oryza sativa nipponbare (GA3) genomic DNA, chromosome 6, BAC
clone:OSUNBa0062J02

Published Only in Database (2001)
2 (bases 1 to 151636)
Sasaki, T., Matsumoto, T. and Yamamoto, K.

Submitted (18-APR-2001) Takuji Sasaki, National Institute of
Agrobiological Sciences, Rice Genome Research Program, Kannondai
2-1-2, Tsukuba, Ibaraki 305-8602, Japan
(E-mail:tsasaki@nias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,
Tel:81-298-38-7441, Fax:81-298-38-7468)

On Aug 24, 2004 this sequence version replaced gi:46849600.
Genes were predicted from the integrated results of the following:
GENSCAN (http://CCR-081.mit.edu/GENSCAN.html), FGENESH

(http://www.softberry.com/), GeneMark.hmm
(http://opal.biology.gatech.edu/GeneMark/), Glimmer-M
(http://www.tigr.org/tdb/glimmer/glmr_form.html), RiceHMM

(http://rgp.dna.affrc.go.jp/RiceHMM/), SplicePredictor
(http://bioinformatics.iastate.edu/cgi-bin/sp.cgi), sim4
(http://globin.cse.psu.edu/html/docs/sim4.html), gap2

(http://www.tigr.org/software/glimmer/), BLASTN and BLASTX. The
genomic sequence was searched against NCBI Nonredundant Protein
database, nr (ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA
sequence database at RGP or DBJ. Protein homologues of the coding
regions were searched against NCBI Nonredundant Protein database
with BLASTP. ESTs represent the identified cDNA sequences using
BLASTN with the corresponding DBJ accession no. and RGP clone ID.
Full-length cDNAs represent the identified cDNA sequences using
BLASTN with the corresponding DBJ accession no.

A gene with identity or significant homology to a protein is
classified based on the protein name to indicate the homology level
such as same name, 'putative-' and '-like protein'. A gene without
significant homology to any protein but with full-length cDNA or
EST homology (covering almost the entire length of partial
sequence) is classified as an 'unknown' protein. A gene predicted
by two or more gene prediction programs is classified as a
'hypothetical' protein according to INSP standard. A gene
predicted by a single gene prediction program is also classified as
a probable 'hypothetical' protein and is included as a
miscellaneous feature of the sequence.

The orientation of the sequence is from M13rev to -21M13 of the BAC
clone. This sequence of OSUNBa0062J02 clone has an overlap with
OJ1226_A12 (DBJ: AP004008) clone at 5' end and with P0453H04

(DBJ: AP005453) at 3' end. Detailed information on overlap and assembly quality together with annotation of this entry is available at <http://xrp.dna.affrc.go.jp/GenomeSeq.html>.

FEATURES		Location/Qualifiers
source	gene	1..151636
		/organism="Oryza sativa (japonica cultivar-group)"
misc_feature	gene	/mol_type="genomic DNA"
		/cultivar="Nipponbare"
		/db_xref="taxon:39947"
		/chromosome="6"
		/clone="OSJNBa0062J02"
		3692..6432
		/gene="OSJNBa0062J02.1"
		join(3692..3890,4145..4201,4709..4952,5371..5483,5920..6006,6088..6432)
		/gene="OSJNBa0062J02.1"
		/note="supported by full-length cDNA(s): AK071267"
CDS	mRNA	join(3834..3890,4145..4201,4709..4952,5371..5483,5920..6006,6088..6207)
		/gene="OSJNBa0062J02.1"
		/note="contains EST(s): D40197 (S1994), AU161639 (S1994)
		contains full-length cDNA(s): AK071267"
		/codon_start=1
		/product="putative ATP synthase delta chain"
		/protein_id="BAD37315.1"
		/db_xref="GI:51535416"
		/translation="MAHLRSLPLLAHLRSLQSGRGPASQVAKPTGDKIKVPEALYGGTGNASALFLTAAKANLDRKVEIRDVVEASKSPLEFSQIKLDSVPKTRVKAI
		TEIFAGSFSDVTKNLAVIDLAVLADRLKHDRIABRFDVLTMAHKGKVKLVIRVILPL
CDS	mRNA	EKEEKELTQDILGKNKLTILIEQIDYSIMGLVIQFGQKVFDMISIKTRAKQEMF
		LRQPLDI"
		complement(7697..10325)
		/gene="OSJNBa0062J02.2"
		complement(join(<7697..7855,7940..8071,8146..8336,8799..8919,9975..>10325))
		/gene="OSJNBa0062J02.2"
		/note="start and end point are not identified"
		complement(join(7697..7855,7940..8071,8146..8336,8799..8919,9975..10325))
		/gene="OSJNBa0062J02.2"
		/note="contains EST(s): C73497 (B4253)"
misc_feature	gene	/codon_start=1
		/product="KNOX family class 2 homeodomain protein"
		/protein_id="BAD37316.1"
		/db_xref="GI:51535417"
		/translation="MAPHYQDHAMDMAAAAEATGGHHHPGVGAGVVGCGGGG
		EREKAAIAAHLPLVLAHVATPVDPQLPRIDAIARPPPLAAATAAAAAA
		ACGAPSGEELDLFMTHYVLLLCFKELQOQVRVHAMEAVMACLEQLTSLTGAS
		PRGSGATMSDDNDVDSSNMFDGNDGDMGFGPLMTEGERSILVERVROELKHE
		LKQYREKLVDIRIELRKRACKLPDGTASTLKNWQASHKWPYPTEEDKARLVQET
		GLQKQINWPFINQRKNWHSNPASSSDSKSKRRRVDF"
misc_feature	gene	join(15289..15348,15401..15532)
		/gene="OSJNBa0062J02.3"
		join(15289..15348,15401..15532)
		/gene="OSJNBa0062J02.3"
		/note="hypothetical ORF
		predicted by GENSCAN
		this category is not included in IRGSP standard"
		complement(17154..23424)
		/gene="OSJNBa0062J02.4"
		complement(join(17154..17767,22937..23424))
CDS	mRNA	/gene="OSJNBa0062J02.4"
		/note="supported by full-length cDNA(s): AK065143"
		complement(join(17744..17767,22937..23424))
		/gene="OSJNBa0062J02.4"
		/note="contains EST(s): AU173291 (R2030)
		contains full-length cDNA(s): AK065143"
		/codon_start=1
		/product="unknown protein"
		/protein_id="BAD37317.1"
		/db_xref="GI:51535418"
CDS	mRNA	/translation="MTAADAEACAAAAADIIICSLRGADLAGWTPPWGTAARKGEVW
		VEEEELAWPTVARGKRSSRRSSPGSGSAATKGRWARGSPASPLDYSGGSGSGS
misc_feature	gene	complement(29040..29397)
		/gene="OSJNBa0062J02.5"
		complement(join(<29040..29235,29324..>29397))
		/gene="OSJNBa0062J02.5"
		/note="start and end point are not identified"
		complement(join(29040..29235,29324..>29397))
		/gene="OSJNBa0062J02.5"
		/note="predicted by GeneMark.hmm etc."
		/codon_start=1
		/product="hypothetical protein"
CDS	mRNA	/protein_id="BAD37318.1"
		/db_xref="GI:51535419"
		/translation="WVATLRRALHLRVLPCPLPFLRRAAALSPTQPPPPPTSSP
		APPSAAGPLGLENTPGMPTRPACNLFLDEMAHKAASYTAMSNRG"
		complement(31053..34828)
		/gene="OSJNBa0062J02.6"
		complement(join(31053..31421,31518..31622,31734..31829,31931..32002,32124..32219,32299..32400,32503..32592,32697..32780,32868..32942,33092..33168,33257..33440,33534..33576,34437..34493,34606..34828))
		/gene="OSJNBa0062J02.6"
		/note="supported by full-length cDNA(s): AK063383"
		complement(join(31380..31421,31518..31622,31734..31829,31931..32002,32124..32219,32299..32400,32503..32592,32697..32780,32868..32942,33092..33142))
CDS	mRNA	/gene="OSJNBa0062J02.6"
		/note="contains full-length cDNA(s): AK063383"
		/codon_start=1
		/product="putative homogenitcic acid geranylgeranyl transferase"
		/protein_id="BAD37319.1"
		/db_xref="GI:51535420"
		/translation="NRSLLDFTMKALMGFEALSSLSCLNNIYVVLNQLYDIQIDKVN
		KPSLPLASGEFSVATCAVLTLTSLMSIAIGRSKAPILCALFIFFLGSAVSVDAP
		LNRKKNFLAASCLIFVRVLVQLAFPAHQHVLKRLPALTYSKVVATLWCCFSS
		VIALFKDIPDIDGRHGFVESLVRGPVWLCINILLITAYGAAILAGASSTNLCO
CDS	mRNA	MITTVFGHGLAFALWQRAQCHDEVENKAWITSFYMFYFWKLYAEYFLIPVQ"
		complement(36006..37518)
		/gene="OSJNBa0062J02.7"
		complement(join(<36006..36167,36607..36675,37426..>37518))
		/gene="OSJNBa0062J02.7"
		/note="start and end point are not identified"
		complement(join(36006..36167,36607..36675,37426..>37518))
		/gene="OSJNBa0062J02.7"
		/note="predicted by GeneMark.hmm etc."
		/codon_start=1
misc_feature	gene	/product="hypothetical protein"
		/protein_id="BAD37320.1"
		/db_xref="GI:51535421"
		/translation="MALTWSRWMLLLRLCTCDQOROKQARMLSAELNGESGSDWA
		RPRETARWEALHIESATVAQPDSPFLSLAYPTARRRRRADEVEGERANLDDELL
		LIFGS"
		complement(37548..37850)
		/gene="OSJNBa0062J02.8"
		complement(37548..37850)
		/gene="OSJNBa0062J02.8"
CDS	mRNA	/note="hypothetical ORF
		predicted by GeneMark.hmm
		this category is not included in IRGSP standard"
		38323..40407
		/gene="OSJNBa0062J02.9"
		join(38323..38611,38695..38811,40088..40407)
		/gene="OSJNBa0062J02.9"
		/note="supported by full-length cDNA(s): AK066837"
		join(38366..38611,38695..38811,40088..40165)
		/gene="OSJNBa0062J02.9"
CDS	mRNA	/note="contains EST(s): AU057515 (S21520), AU057516 (S21520)
		contains full-length cDNA(s): AK066837"
		/codon_start=1
		/product="putative 50S ribosomal protein L35"
		77.6%; Score 19.4; DB 15; Length 151636;

Best Local Similarity 95.2%; Pred. No. 2.9e+02; Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 AAAACAATGACCAACTGGGCT 23
|||||
Db 131918 AAAATAATGACCAACTGGGCT 131898

RESULT 37
AP005453/c
LOCUS
DEFINITION
Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 6,
PAC clone:P0453H04.
AP005453
AP005453.1 GI:21624005

ORyza sativa (japonica cultivar-group)
ORyza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.

1
Sasaki, T., Matsumoto, T. and Katayose, Y.
Oryza sativa nipponbare(GA3) genomic DNA, chromosome 6, PAC
clone:P0453H04
Published Only in Database (2002)
2 (bases 1 to 175047)
Sasaki, T., Matsumoto, T. and Katayose, Y.
Direct Submission
Submitted (20-JUN-2002) Takuji Sasaki, National Institute of
Agrbiological Sciences, Rice Genome Research Program; Kannondai
2-1-2, Tsukuba, Ibaraki 305-8602, Japan
(E-mail:tsasaki@nias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,
Tel:81-298-38-7441, Fax:81-298-38-7469)
Genes were predicted from the integrated results of the following:
GENSCAN (http://CCR-081.mit.edu/GENSCAN.html), FGENESH
(http://www.softberry.com/), GeneMark hmh
(http://opal.biology.gatech.edu/GeneMark/), GlimmerM
(http://www.tigr.org/tdb/glimmer/glmr_form.html), RiceHMM
(http://rgp.dna.affrc.go.jp/RiceHMM/), SplicePredictor
(http://bioinformatics.tastate.edu/cgi-bin/sp.cgi), sim4
(http://globin.cse.psu.edu/html/docs/sim4.html), gap2
(http://www.tigr.org/software/glimmer/), BLASTN and BLASTX. The
genomic sequence was searched against NCBI NonRedundant Protein
database, nr (ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA
sequence database at RGP or DDBJ. Protein homologs of the coding
regions were searched against NCBI NonRedundant Protein database
with BLASTP. ESTs represent the identified cDNA sequences using
BLASTN with the corresponding DDBJ accession no. and RGP clone ID.
Full-length cDNAs represent the identified cDNA sequences using
BLASTN with the corresponding DDBJ accession no.

A gene with identity or significant homology to a protein is
classified based on the protein name to indicate the homology level
such as same name, 'putative-' and '-like protein'. A gene without
significant homology to any protein but with full-length cDNA or
EST homology (covering almost the entire length of partial
sequence) is classified as an 'unknown' protein. A gene predicted
by two or more gene prediction programs is classified as a
'hypothetical' protein according to IRGSP standard. A gene
predicted by a single gene prediction program is also classified as a
probable 'hypothetical' protein and is included as a
miscellaneous feature of the sequence.
The orientation of the sequence is from 3p6 to T7 of the PAC clone.
This sequence of P0453H04 clone has an overlap with OSUNBA0062J02
(DDBJ: AP003517) clone at 5' end and with OSUNBA0085J13 (DDBJ:
AP003565) at 3' end. Detailed information on overlap and assembly
quality together with annotation of this entry is available at
http://rgp.dna.affrc.go.jp/genomeSeq.html.

FEATURES
source
1. .175047
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"

/db_xref="taxon:39947"
/chromosome="6"
/clone="P0453H04"
complement(2482..2691)
/gene="P0453H04.1"
complement(2482..2691)
/gene="P0453H04.1"
/note="hypothetical ORF
predicted by glimmerM
this category is not included in IRGSP standard"
complement(join(3217..3678,5673..5759,5979..6023))
/gene="P0453H04.2"
complement(join(3217..3678,5673..5759,5979..6023))
/gene="P0453H04.2"
/note="hypothetical ORF
predicted by GeneMark.hmm
this category is not included in IRGSP standard"
join(7151..7301,7947..8340,8427..8556)
/gene="P0453H04.3"
join(7151..7301,7947..8340,8427..8556)
/gene="P0453H04.3"
/note="hypothetical ORF
predicted by GENSCAN
this category is not included in IRGSP standard"
complement(8579..9350)
/gene="P0453H04.4"
complement(join(<8579..8890,9342..>9350))
/note="start and end point are not identified"
complement(join(8579..8890,9342..9350))
/gene="P0453H04.4"
/note="predicted by GENSCAN etc."
/codon_start=1
/product="hypothetical protein"
/protein_id="BAD38098.1"
/db_xref="GI:51534974"
/translation="MTTQKDEQQRPAAGGSRSDGGRWAEGRAAGRGSGRAAGGGE
MTGGRRDENAAREEASSSRPVEEAAVQREGGAAVAVGGGPMGGRKSGSGSP
TSRP"
complement(9530..9809)
/gene="P0453H04.5"
complement(join(<9530..9605,9700..>9809))
/gene="P0453H04.5"
/note="start and end point are not identified"
complement(join(9530..9605,9700..9809))
/gene="P0453H04.5"
/note="predicted by GeneMark.hmm etc."
/codon_start=1
/product="hypothetical protein"
/protein_id="BAD38099.1"
/db_xref="GI:51534975"
/translation="MTHVLCSMFLSVAAPFDGSFEAAKVLAFVAGGDGLMGAVAVAS
HEGADPSGQAGRPRA"
10192..18908
/gene="P0453H04.6"
join(10192..10419,11280..11361,11889..11967,12067..12164,
12292..12348,12932..13019,13129..13238,13421..13522,
13777..13928,14729..14839,14915..14984,15075..15114,
15528..15692,15794..15870,17190..17306,17445..17710,
18230..18325,18409..18475,18568..18908)
/gene="P0453H04.6"
/note="supported by full-length cDNA(s): AK072591"
join(10286..10419,11280..11361,11889..11967,12067..12164,
12292..12348,12932..13019,13129..13238,13421..13522,
13777..13928,14729..14839,14915..14984,15075..15114,
15528..15692,15794..15870,17190..17306,17445..17710,
18230..18325,18409..18475,18568..18621)
/gene="P0453H04.6"
/note="TATA binding protein associated factor 5
contains full-length cDNA(s): AK072591"
/codon_start=1
/product="putative TAF5"
/protein_id="BAD38100.1"

```
/db_xref="GI:51534976"
/translation="MEDEMEKKVQOYLQKGFRLTELALQERNRISTSVSDVALA
RENDPARYDGYSLRTWAYSSLDQYKHELLRLVLPVFIHSFMDLVAEGHTQEARSE
FHTFHEDHEMSRLQKLEGLISPSHLEMLARSLRNQRFRIKCEYSYELLQOYL
QKQALVLVIGIINERTFFVSPQPSLISDDTDVVALVGTKKDLAKVINKLKEVHGLL
EDSVEMRKTLLESDEKTEAEKDAEDNNRKSSEGGQGVKVKKKDKIAGATG
KTKNSSTSVSHAPRVKPELTLPVIEVEQSLIEDLRNRAOLNSIALPSVSYFTFLN
TINGLNCSSISHDGSLVGGFSDSVKVMDSKIGOPPKTSSPOGBGLSQGERTSAS
DVGKRYPTILFQHGSGRVYSARPFGDFLLSSSDSTIELWSTKLNANLVCYKGNYP
VMDVQSPVGHVTFASHDRKTARIWMDKIQPIRMAGHLSVDVDCVQHVNCVNIATG
SDKTVLMDVDTGECIRFIHRSNVLKSLAWSPOGRYMASGDEGTINMDLSSGRK
VSPLGHSVCVMSLAYSCGALIASGADCTVKLDVNASSTKVLKTDSTNRLRLMK
TLTKGTPTVTLRFSRNLFAAGALSIGS"
complement (20251..20484)
/gene="P0453H04.7"
complement (<20251..>20484)
/genes="P0453H04.7"
/notes="start and end point are not identified"
complement (20251..20484)
/gene="P0453H04.7"
/notes="predicted by GlimmerM etc."
/codon_start=1
/product="hypothetical protein"
/protein_id="BAD38101.1"
/db_xref="GI:51534977"
/translation="MFGPHVVAGISIGQAGTGGAVRQGRRCALSPGSGQAGVDASG
ARQGGQDPCDRGELGRCWFSAIGLVVAGN"
complement (join(21526..21633,22311..22403,25381..25499,
28308..28323,28773..28832))
/genes="P0453H04.8"
complement (join(21526..21633,22311..22403,25381..25499,
28308..28323,28773..28832))
/genes="P0453H04.8"
/notes="hypothetical ORF
predicted by GENSCAN
this category is not included in IRGSP standard"
32254..32421
/gene="P0453H04.9"
32254..32421
/genes="P0453H04.9"
/notes="hypothetical ORF
predicted by GlimmerM
this category is not included in IRGSP standard"
35423..41646
/genes="P0453H04.10"
join(<35423..36233,36603..36722,37003..37116,41025..41143,
41257..>41646)
/genes="P0453H04.10"
/notes="supported by full-length cDNA(s): AK103769"
35423..41646
/gene="P0453H04.10"
/notes="contains full-length cDNA(s): AK103769
non-coding transcript
probably inactive due to including stop codon(s) in CDS"
complement (join(41799..41902,44014..44110))
/gene="P0453H04.11"
complement (join(41799..41902,44014..44110))
/genes="P0453H04.11"
/notes="hypothetical ORF
predicted by RiceHMM
this category is not included in IRGSP standard"
44222..46867
/gene="P0453H04.12"
<44222..>46867

Query Match 77.6%; Score 19.4; DB 15; Length 175047;
Best Local Similarity 95.2%; Pred. No. 2.9e+02;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 AAAACAATGACCAACTGGGCT 23
Db 17213 AAAATAATGACCAACTGGGCT 17193
```

```
RESULT 38
CQ337709
LOCUS CQ337709 600 bp DNA linear PAT 23-JAN-2004
DEFINITION Sequence 11803 from Patent WO0157275.
ACCESSION CQ337709
VERSION CQ337709.1 GI:41286780
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
REFERENCE
1 Penn, S.G., Hanzel, D.K., Chen, W. and Rank, D.R.
Human genome-derived single exon nucleic acid probes useful for
analysis of gene expression in human brain
Patent: WO 0157275-A 11803 09-AUG-2001;
Aeomica, Inc. (US)
FEATURES
Location/Qualifiers
1..600
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="MAP TO AC009178.4
EXPRESSED IN BRAIN, SIGNAL = 1.3"
ORIGIN
Query Match 76.8%; Score 19.2; DB 6; Length 600;
Best Local Similarity 87.5%; Pred. No. 4.1e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 AAAACAATGACCAACTGGGCTGT 25
Db 370 AAAACAATGACCAACTGGTTGT 393

RESULT 39
AK102491
LOCUS AK102491 1310 bp mRNA linear PLN 24-JUL-2003
DEFINITION Oryza sativa (japonica cultivar-group) cDNA clone:J033094019, full
insert sequence.
ACCESSION AK102491
VERSION AK102491.1 GI:32987700
KEYWORDS FLI_CDNA; CAP trapper.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE
1 The Rice Full-length cDNA Consortium, National Institute of
Agrobiological Sciences Rice Full-length cDNA Project team,
Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K.,
Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I.,
Kojima, K., Nami, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C.,
Ohtsuki, K., Shishiki, T., Foundation of Advancement of International
Science Genome Sequencing & Analysis Group, Otono, Y., Murakami, K.,
Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y.,
Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M.,
Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J.,
Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J.,
Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN;
Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S.,
Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M.,
Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Osato, N., Ota, Y.,
Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,
Yoshino, M. and Hayashizaki, Y.
Collection, mapping, and annotation of over 28,000 cDNA clones from
Japonica rice
Science 301 (5631), 376-379 (2003)
REFERENCE
1 12869764
2 (bases 1 to 1310)
```

AUTHORS Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Doi,K., Fujimura,T., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayashizaki,Y., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F., Hotta,I., Iida,J., Iida,Y., Ikeda,R., Imamura,K., Imotoani,K., Ishibiki,J., Ishii,Y., Ishikawa,M., Itoh,M., Kagawa,I., Kanagawa,S., Kato,H., Kawagashira,N., Kawai,J., Kawamata,M., Kikuchi,S., Kishikawa-Hirozane,T., Kishimoto,N., Kobayashi,M., Kodama,T., Kojima,K., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Kuroaki,T., Kusumegi,T., Li,C., Lu,M., Masuda,H., Matsubara,K., Matsuyama,T., Miura,J., Miyazaki,A., Mizuno,K., Murakami,K., Murata,M., Nagata,T., Nakamura,M., Namiki,T., Nariikawa,R., Nikura,J., Nishi,K., Nomura,K., Numasaki,R., Ohneda,E., Ohno,M., Ohtsuki,K., Oka,M., Ooka,H., Osato,N., Ota,Y., Otomo,Y., Ryu,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Sato,K., Satoh,K., Shibata,K., Shinagawa,A., Shiraki,T., Shishiki,T., Sogabe,Y., Sugano,S., Sugiyama,A., Suzuki,K., Suzuki,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Tsunoda,Y., Ueda,M., Waki,K., Xie,Q., Yahagi,W., Yamada,H., Yamamoto,M., Yasunishi,A., Yazaki,J., Yokomizo,S. and Yoshimura,A.

TITLE Direct Submission

JOURNAL Submitted (27-AUG-2002) Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail:skikuchi@ias.affrc.go.jp, Tel:81-29-838-7007, Fax:81-29-838-7007)

COMMENT This clone is one of the 28K full-length cDNA clones from japonica rice. URL : <http://cdna01.dna.affrc.go.jp/cDNA/> NIAS Rice Full-Length cDNA Project Team: Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T. and Yamamoto,M.

FAIS Genome Sequencing & Analysis Group: Otomo,Y., Iida,Y., Fujimura,T., Ikeda,R., Ishibiki,J., Kawamata,M., Kobayashi,M., Kodama,T., Kuroaki,T., Kusumegi,T., Lu,M., Masuda,H., Miura,J., Mizuno,K., Nariikawa,R., Nikura,J., Oka,M., Ryu,R., Sugano,S., Sugiyama,A., Suzuki,Y., Tsunoda,Y., Ueda,M., Xie,Q., Yokomizo,S., Yoshimura,A., Matsubara,K. and Murakami,K.

Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F., Iida,J., Imamura,K., Imotoani,K., Ishii,Y., Itoh,M., Kagawa,I., Kanagawa,S., Kato,H., Kawai,J., Kishikawa-Hirozane,T., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Ohno,M., Osato,N., Ota,Y., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Waki,K., Yasunishi,A. and Hayashizaki,Y.

FEATURES Location/Qualifiers

1. .1310

source

location=Oryza sativa (japonica cultivar-group)"

/organism="Oryza sativa (japonica cultivar-group)"

/mol_type="mRNA"

/cultivar="Nipponbare"

/db_xref="taxon:39947"

/clone="J033094019"

ORIGIN

Query Match 76.8%; Score 19.2; DB 15; Length 1310;

Best Local Similarity 87.5%; Pred. No. 4e+02;

Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAAAATGACCAACTGGGCTGT 25

|||||

DB 1229 AAAAAAGTGTGACCAACTGGGCTGT 1252

|||||

RESULT 40

LOCUS AP004334

DEFINITION Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 7, PAC clone:P0455H11.

ACCESSION AP004334

VERSION AP004334.2 GI:21203158

KEYWORDS

SOURCE Oryza sativa (japonica cultivar-group)

ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.

1

Sasaki,T., Matsumoto,T. and Yamamoto,K.

Oryza sativa nipponbare (G3) genomic DNA, chromosome 7, PAC clone:P0455H11

Published Only in Database (2001)

2 (bases 1 to 101607)

Sasaki,T., Matsumoto,T. and Yamamoto,K.

Direct Submission

Submitted (14-NOV-2001) Takuji Sasaki, National Institute of Agrobiological Sciences, Rice Genome Research Program; Kannondai 2-1-2, Tsukuba, Ibaraki, 305-8602, Japan (E-mail:tsasaki@ias.affrc.go.jp, URL:<http://rgp.dna.affrc.go.jp/>, Tel:81-298-38-7441, Fax:81-298-38-7468)

On May 24, 2002 this sequence version replaced gi:16930119.

Genes were predicted from the integrated results of the following: GENSCAN (<http://CCR-081.mit.edu/GENSCAN.html>), GENESH (<http://www.softberry.com/>), GeneMark.hmm (<http://opal.biology.gatech.edu/GeneMark/>), GlimmerM (<http://opal.tigr.org/tdb/glimmerm/glmr form.html>), RiceHMM (<http://rgp.dna.affrc.go.jp/RiceHMM/>), SplicePredictor (<http://bioinformatics.iascate.edu/cgi-bin/sp.cgi>), sim4 (<http://globin.cse.psu.edu/html/docs/sim4.html>), gap2 (<http://www.tigr.org/software/glimmer/>), BLASTN and BLASTX. The genomic sequence was searched against NCBI NonRedundant Protein database, nr (<ftp://ncbi.nlm.nih.gov/blast/db>) and the cDNA sequence database at RGP or DDBJ. Protein homologies of the coding regions were searched against NCBI NonRedundant Protein database with BLASTP. ESTs represent the identified cDNA sequences using BLASTN with the corresponding DDBJ accession no. and RGP clone ID. Full-length cDNAs represent the identified cDNA sequences using BLASTN with the corresponding DDBJ accession no.

A gene with identity or significant homology to a protein is classified based on the protein name to indicate the homology level such as same name, 'putative-' and '-like protein'. A gene without significant homology to any protein but with full-length cDNA or EST homology (covering almost the entire length of partial sequence) is classified as an 'unknown' protein. A gene predicted by two or more gene prediction programs is classified as a 'hypothetical' protein according to IRGSP standard. A gene predicted by a single gene prediction program is also classified as a probable 'hypothetical' protein and is included as a miscellaneous feature of the sequence.

The orientation of the sequence is from T7 to SP6 of the PAC clone. This orientation of P0455H11 clone has an overlap with P0519E12 clone (DDBJ: AP004339) at 5' end and an overlap with OJ1332.C12 clone (DDBJ: AP003752) at 3' end. Detailed information on overlap and assembly quality together with annotation of this entry is available at <http://rgp.dna.affrc.go.jp/GenomeSeq.html>.

Location/Qualifiers

1. .101607

organism="Oryza sativa (japonica cultivar-group)"

/mol_type="genomic DNA"

/cultivar="Nipponbare"

/db_xref="taxon:39947"

/chromosome="7"

/clone="P0455H11"

Join(4486. .4493,5998. .6098,7842. .7909)

/gene="P0455H11.101"

Join(4486. .4493,5998. .6098,7842. .7909)

misc_feature

FEATURES source

gene

misc_feature

gap /estimated_length=unknown
20776..20875
gap /estimated_length=unknown
33308..33407
gap /estimated_length=unknown
43459..43558
gap /estimated_length=unknown
58213..58312
gap /estimated_length=unknown
73363..73462
gap /estimated_length=unknown
85518..85617
gap /estimated_length=unknown
98209..98308
gap /estimated_length=unknown
115205..115304
gap /estimated_length=unknown
131266..131365
gap /estimated_length=unknown
160956..161055
gap /estimated_length=unknown
202836..202935
gap /estimated_length=unknown
247059..247158
gap /estimated_length=unknown
292711..292810
gap /estimated_length=unknown

ORIGIN

Query Match 76.8%; Score 19.2; DB 14; Length 110000;
Best Local Similarity 87.5%; Pred. No. 3.7e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AAAAAACAATGACCAACTGGGCTG 24
||||||| |

Db 98079 AAAAAACAACAGCAACTGGGCTG 98102

RESULT 42
AP008213_262
WPCOMMENT

Sequence split into 297 fragments LOCUS AP008213 Accession AP008213

Fragment Name	Begin	End
AP008213_000	1	110000
AP008213_001	100001	210000
AP008213_002	200001	310000
AP008213_003	300001	410000
AP008213_004	400001	510000
AP008213_005	500001	610000
AP008213_006	600001	710000
AP008213_007	700001	810000
AP008213_008	800001	910000
AP008213_009	900001	1010000
AP008213_010	1000001	1110000
AP008213_011	1100001	1210000
AP008213_012	1200001	1310000
AP008213_013	1300001	1410000
AP008213_014	1400001	1510000
AP008213_015	1500001	1610000
AP008213_016	1600001	1710000
AP008213_017	1700001	1810000
AP008213_018	1800001	1910000
AP008213_019	1900001	2010000
AP008213_020	2000001	2110000
AP008213_021	2100001	2210000
AP008213_022	2200001	2310000
AP008213_023	2300001	2410000
AP008213_024	2400001	2510000
AP008213_025	2500001	2610000
AP008213_026	2600001	2710000
AP008213_027	2700001	2810000
AP008213_028	2800001	2910000
AP008213_029	2900001	3010000

AP008213_030	3000001	3110000
AP008213_031	3100001	3210000
AP008213_032	3200001	3310000
AP008213_033	3300001	3410000
AP008213_034	3400001	3510000
AP008213_035	3500001	3610000
AP008213_036	3600001	3710000
AP008213_037	3700001	3810000
AP008213_038	3800001	3910000
AP008213_039	3900001	4010000
AP008213_040	4000001	4110000
AP008213_041	4100001	4210000
AP008213_042	4200001	4310000
AP008213_043	4300001	4410000
AP008213_044	4400001	4510000
AP008213_045	4500001	4610000
AP008213_046	4600001	4710000
AP008213_047	4700001	4810000
AP008213_048	4800001	4910000
AP008213_049	4900001	5010000
AP008213_050	5000001	5110000
AP008213_051	5100001	5210000
AP008213_052	5200001	5310000
AP008213_053	5300001	5410000
AP008213_054	5400001	5510000
AP008213_055	5500001	5610000
AP008213_056	5600001	5710000
AP008213_057	5700001	5810000
AP008213_058	5800001	5910000
AP008213_059	5900001	6010000
AP008213_060	6000001	6110000
AP008213_061	6100001	6210000
AP008213_062	6200001	6310000
AP008213_063	6300001	6410000
AP008213_064	6400001	6510000
AP008213_065	6500001	6610000
AP008213_066	6600001	6710000
AP008213_067	6700001	6810000
AP008213_068	6800001	6910000
AP008213_069	6900001	7010000
AP008213_070	7000001	7110000
AP008213_071	7100001	7210000
AP008213_072	7200001	7310000
AP008213_073	7300001	7410000
AP008213_074	7400001	7510000
AP008213_075	7500001	7610000
AP008213_076	7600001	7710000
AP008213_077	7700001	7810000
AP008213_078	7800001	7910000
AP008213_079	7900001	8010000
AP008213_080	8000001	8110000
AP008213_081	8100001	8210000
AP008213_082	8200001	8310000
AP008213_083	8300001	8410000
AP008213_084	8400001	8510000
AP008213_085	8500001	8610000
AP008213_086	8600001	8710000
AP008213_087	8700001	8810000
AP008213_088	8800001	8910000
AP008213_089	8900001	9010000
AP008213_090	9000001	9110000
AP008213_091	9100001	9210000
AP008213_092	9200001	9310000
AP008213_093	9300001	9410000
AP008213_094	9400001	9510000
AP008213_095	9500001	9610000
AP008213_096	9600001	9710000
AP008213_097	9700001	9810000
AP008213_098	9800001	9910000
AP008213_099	9900001	10010000
AP008213_100	10000001	10110000
AP008213_101	10100001	10210000
AP008213_102	10200001	10310000

AP008213_103 10300001 10410000
 AP008213_104 10400001 10510000
 AP008213_105 10500001 10600000
 AP008213_106 10600001 10710000
 AP008213_107 10700001 10810000
 AP008213_108 10800001 10910000
 AP008213_109 10900001 11010000
 AP008213_110 11000001 11110000
 AP008213_111 11100001 11210000
 AP008213_112 11200001 11310000
 AP008213_113 11300001 11410000
 AP008213_114 11400001 11510000
 AP008213_115 11500001 11610000
 AP008213_116 11600001 11710000
 AP008213_117 11700001 11810000
 AP008213_118 11800001 11910000
 AP008213_119 11900001 12010000
 AP008213_120 12000001 12110000
 AP008213_121 12100001 12210000
 AP008213_122 12200001 12310000
 AP008213_123 12300001 12410000
 AP008213_124 12400001 12510000
 AP008213_125 12500001 12610000
 AP008213_126 12600001 12710000
 AP008213_127 12700001 12810000
 AP008213_128 12800001 12910000
 AP008213_129 12900001 13010000
 AP008213_130 13000001 13110000
 AP008213_131 13100001 13210000
 AP008213_132 13200001 13310000
 AP008213_133 13300001 13410000
 AP008213_134 13400001 13510000
 AP008213_135 13500001 13610000
 AP008213_136 13600001 13710000
 AP008213_137 13700001 13810000
 AP008213_138 13800001 13910000
 AP008213_139 13900001 14010000
 AP008213_140 14000001 14110000
 AP008213_141 14100001 14210000
 AP008213_142 14200001 14310000
 AP008213_143 14300001 14410000
 AP008213_144 14400001 14510000
 AP008213_145 14500001 14610000
 AP008213_146 14600001 14710000
 AP008213_147 14700001 14810000
 AP008213_148 14800001 14910000
 AP008213_149 14900001 15010000
 AP008213_150 15000001 15110000
 AP008213_151 15100001 15210000
 AP008213_152 15200001 15310000
 AP008213_153 15300001 15410000
 AP008213_154 15400001 15510000
 AP008213_155 15500001 15610000
 AP008213_156 15600001 15710000
 AP008213_157 15700001 15810000
 AP008213_158 15800001 15910000
 AP008213_159 15900001 16010000
 AP008213_160 16000001 16110000
 AP008213_161 16100001 16210000
 AP008213_162 16200001 16310000
 AP008213_163 16300001 16410000
 AP008213_164 16400001 16510000
 AP008213_165 16500001 16610000
 AP008213_166 16600001 16710000
 AP008213_167 16700001 16810000
 AP008213_168 16800001 16910000
 AP008213_169 16900001 17010000
 AP008213_170 17000001 17110000
 AP008213_171 17100001 17210000
 AP008213_172 17200001 17310000
 AP008213_173 17300001 17410000
 AP008213_174 17400001 17510000
 AP008213_175 17500001 17610000

AP008213_176 17600001 17710000
 AP008213_177 17700001 17810000
 AP008213_178 17800001 17910000
 AP008213_179 17900001 18010000
 AP008213_180 18000001 18110000
 AP008213_181 18100001 18210000
 AP008213_182 18200001 18310000
 AP008213_183 18300001 18410000
 AP008213_184 18400001 18510000
 AP008213_185 18500001 18610000
 AP008213_186 18600001 18710000
 AP008213_187 18700001 18810000
 AP008213_188 18800001 18910000
 AP008213_189 18900001 19010000
 AP008213_190 19000001 19110000
 AP008213_191 19100001 19210000
 AP008213_192 19200001 19310000
 AP008213_193 19300001 19410000
 AP008213_194 19400001 19510000
 AP008213_195 19500001 19610000
 AP008213_196 19600001 19710000

Query Match 76.8%; Score 19.2; DB 15; Length 110000;
 Best Local Similarity 87.5%; Pred. NO. 3.7e+02;
 Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAACAATGACCACTGGGCTGT 25
 DB 70459 AAAAAGTGTGACCAACTGGGCTGT 70482

RESULT 43
 AC152963

LOCUS AC152963 111804 bp DNA linear HTG 03-DEC-2004
 DEFINITION Takifugu rubripes clone MRC-195B17, WORKING DRAFT SEQUENCE.

ACCESSION AC152963

VERSION AC152963.1 GI:56089309

KEYWORDS HTG; HTGS PHASE2; HTGS DRAFT.

SOURCE Takifugu rubripes (Fugu rubripes)

ORGANISM

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
 Tetraodontoidea; Tetraodontidae; Takifugu.

REFERENCE 1 (bases 1 to 111804)
 AUTHORS Cheng, J.-F., Hamilton, M., Peng, Y., Hosseini, R., Peng, Z., Malinov, I. and Rubin, E.M.

TITLE Direct Submission

REFERENCE 2 (bases 1 to 111804)

AUTHORS Cheng, J.-F., Hamilton, M., Peng, Y., Hosseini, R., Peng, Z., Malinov, I. and Rubin, E.M.

TITLE Direct Submission

JOURNAL Submitted (25-NOV-2004) Genome Sciences, Lawrence Berkeley National Laboratory, 1 Cyclotron Rd., Berkeley, CA 94720, USA

REFERENCE 3 (bases 1 to 111804)

AUTHORS Cheng, J.-F., Hamilton, M., Peng, Y., Hosseini, R., Peng, Z., Malinov, I. and Rubin, E.M.

TITLE Direct Submission

JOURNAL Submitted (03-DEC-2004) Genome Sciences, Lawrence Berkeley National Laboratory, 1 Cyclotron Rd., Berkeley, CA 94720, USA

COMMENT

Sequence Produced by Berkeley PGA

Web site: <http://pga.lbl.gov>

Center Code: PGABERK

Center Project Name: F072

Bac Clone Name: MRC-195B17

This sequence has been compared to sequences of other species

using Vista (<http://www-gsd.lbl.gov/VISTA>). The results can be

viewed at:
http://pga.lbl.gov/cgi-bin/search_cvcgd?type=n&value=FTK7

The order-orientation of the draft sequence was accomplished by using:
Avid (<http://baboon.math.berkeley.edu/mavid>) and paired end information.

Funding agent: Programs for Genomic Applications (NHLBI)

Summary Statistics:

Sequencing vector: Plasmid; pUC18

Chemistry: Dye-terminator Big Dye

Assembly program: Phrap version 0.990329.

* NOTE: This is a 'working draft' sequence. It currently

* consists of 1 contigs. Gaps between the contigs

* are represented as runs of N. The order of the pieces

* is believed to be correct as given, however the sizes

* of the gaps between them are based on estimates that have

* provided by the submitter.

* This sequence will be replaced

* by the finished sequence as soon as it is available and

* the accession number will be preserved.

* 1 111804: contig of 111804 bp in length.

Location/Qualifiers

1. 111804

/organism="Takifugu rubripes"

/mol_type="genomic DNA"

/db_xref="taxon:31033"

/clone="MRC-195B17"

ORIGIN

Query Match 76.8%; Score 19.2; DB 14; Length 111804;
Best Local Similarity 87.5%; Pred. NO. 3.7e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AAAAACAATGACCACTGGCGTG 24

Db 81336 AAAGAACAATGACCACTGGCGTG 81359

RESULT 44

AC117517/c

LOCUS

DEFINITION Homo sapiens X BAC RP13-356M24 (Roswell Park Cancer Institute Human BAC Library) complete sequence.

AC117517

AC117517.7 GI:21263143

HTG.

Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

1 (bases 1 to 121628)

Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-Osman,F.R., Allen,C., Albrooks,S.L., Amarantunge,H.C., Are,J.R., Banks,T., Barbaria,J., Benton,J., Bivaga,K., Blankenburg,K., Bonnin,D., Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F., Carter,M., Cavazos,S., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H., Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hoques,M., Holloway,C., Hollins,B., Homs,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,K., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,

Louisegeed,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E., Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenwo,S., Ogutu,H., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.D., Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shoohtari,N., Sleson,I., Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R., Wang,Q., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S., Williams,G., Williamson,A., Wleczkyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D., Weinstock,G. and Gibbs,R.

TITLE

Direct Submission

Unpublished

REFERENCE 2 (bases 1 to 121628)

Worley,K.C.

Direct Submission

Submitted (10-APR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE 3 (bases 1 to 121628)

Worley,K.C.

Direct Submission

Submitted (26-MAY-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE 4 (bases 1 to 121628)

Worley,K.C.

Direct Submission

Submitted (30-MAY-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE 5 (bases 1 to 121628)

Worley,K.C.

Direct Submission

Submitted (24-OCT-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE 6 (bases 1 to 121628)

Worley,K.C.

Direct Submission

Submitted (15-MAR-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On May 30, 2002 this sequence version replaced gi:21217391.

INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email

gc-help@bcm.tmc.edu

COMMENT

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.

ANNOTATION OF FEATURES:

STSs are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts

Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences.

Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL: <http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

QUALSTAT-REPORT.

```

FEATURES
  source      Location/Qualifiers
              1..121628
               /organism="Homo sapiens"
               /mol_type="genomic DNA"
               /db_xref="taxon:9606"
               /chromosome="X"
  misc_feature /clone="RP13-356M24"
               complement(1..2048)
               /notes="overlaps bases 2..2049 of clone AC004673"
               /functions="clone overlap"
               complement(10..291)
               /rpt_family="L1"
  repeat_region 283..361
               /rpt_family="L1MA10"
  repeat_region 362..395
               /rpt_family="CA)n"
  repeat_region 396..548
               /rpt_family="L1MA10"
  repeat_region 744..1126
               /rpt_family="WSTD"
               complement(1573..1792)
               /rpt_family="L2"
  repeat_region complement(2105..2165)
               /rpt_family="MERSA"
  repeat_region 2547..2786
               /rpt_family="LTR16C"
  repeat_region 2993..3062
               /rpt_family="L2"
  repeat_region 3080..3512
               /rpt_family="L2"
  repeat_region complement(3513..3801)
               /rpt_family="AluY"
  repeat_region 3802..4211
               /rpt_family="L2"
  repeat_region complement(4209..4323)
               /rpt_family="L2"
  STS           6708..6819
               /standard_name="47416"
  STS           6758..6925
               /standard_name="42458"
  STS           6763..6872
               /standard_name="84773"
               complement(6993..7308)
               /rpt_family="AluYb8"
  repeat_region complement(8308..8569)
               /rpt_family="L2"
  repeat_region 8704..8739
               /rpt_family="T-rich"
  repeat_region complement(8740..9051)
               /rpt_family="AluX"
  repeat_region complement(9062..10864)
               /rpt_family="L1PA16"
  repeat_region 10870..10986
               /rpt_family="AluDb"
  repeat_region 10987..11017
               /rpt_family="TANA)n"
  repeat_region complement(11028..12468)

```

```

repeat_region /rpt_family="L1PA16"
               complement(12469..12777)
               /rpt_family="AluY"
repeat_region 12778..12905
               /rpt_family="MER75"
               complement(12922..13566)
               /rpt_family="L1PA16"
repeat_region 13567..14142
               /rpt_family="MLT2D"
               complement(14143..14532)
               /rpt_family="L1PA16"
repeat_region 14533..15022
               /rpt_family="MER75"
               complement(15051..15434)
               /rpt_family="L1PA16"
repeat_region complement(15395..15772)
               /rpt_family="L1PA13"
repeat_region complement(15779..16579)
               /rpt_family="L1PA13"
repeat_region 16770..16969
               /rpt_family="MIR"

```

Query Match 76.8%; Score 19.2; DB 8; Length 121628;
 Best Local Similarity 87.5%; Pred. No. 3.7e+02;
 Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

Qy 2 AAAACAATGACCAACTGGCTGT 25
Db 3517 AAAACAATGATAAAGTGGACTGT 3494

```

RESULT 45

```

AC144797      129045 bp DNA linear ROD 15-MAY-2004
LOCUS Mus musculus BAC clone RP24-349B2 from chromosome 16, complete
DEFINITION sequence.
ACCESSION AC144797.2 GI:45237305
VERSION AC144797.2 GI:45237305
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
           Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
           Sciurognathi; Muridea; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 129045)
          VanBrunt,A., Van Brunt,A., Bielicki,L., Meyer,R. and Haakenson,W.
          The sequence of Mus musculus BAC clone RP24-349B2
          Unpublished (2001)
REFERENCE 2 (bases 1 to 129045)
          Wilson,R.K.
          Direct Submission
          Submitted (20-MAY-2003) Genome Sequencing Center, 4444 Forest Park
          Parkway, St. Louis, MO 63108, USA
REFERENCE 3 (bases 1 to 129045)
          Wilson,R.K.
          Direct Submission
          Submitted (06-MAR-2004) Genome Sequencing Center, 4444 Forest Park
          Parkway, St. Louis, MO 63108, USA
REFERENCE 4 (bases 1 to 129045)
          Wilson,R.K.
          Direct Submission
          Submitted (15-MAY-2004) Genome Sequencing Center, 4444 Forest Park
          Parkway, St. Louis, MO 63108, USA
          On Mar 6, 2004 this sequence version replaced gi.30911150.
          ----- Genome Center
          Center: Washington University Genome Sequencing Center
          Center code: WUGSC
          Web site: http://genome.wustl.edu
          Contact: submissions@wustl.edu
          ----- Summary Statistics
          -----
          Center project name: M_BB0349B02
          -----

```

NOTICE:

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone, fosmid clone or direct clone walk sequence. Sequence from the Mouse Genome Sequencing Consortium whole genome shotgun may have been used to obtain the consensus sequence; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. Wes Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu>

SOURCE INFORMATION:

The RPCI-24 BAC Library has been constructed by Pieter de Jong and coworkers (<http://www.chori.org>) from male C57BL/6J mouse spleen and/or brain genomic DNA. The clone and detailed information can be obtained from Pieter de Jong and coworkers at <http://www.chori.org>

NEIGHBORING SEQUENCE INFORMATION:

This sequence is the entire insert of the clone.

FEATURES

```

source
    Location/Qualifiers
      1..129045
        /organism="Mus musculus"
        /mol_type="genomic DNA"
        /db_xref="taxon:10090"
        /chromosome="16"
        /map="16"
        /clone="RP24-349B2"
        /clone_lib="RPCI-24"
        /size="28380"
        /product="tRNA-Ser"
        /note="Likely pseudogene (HMM Sc=33.19 / Sec struct
          43749..43816
          /product="tRNA-Ser"
          /notes="Likely pseudogene (HMM Sc=18.82 / Sec struct
            Sc=1.22)"
          /product="tRNA-Ser"
          /notes="Likely pseudogene (HMM Sc=33.19 / Sec struct
            Sc=10.52)"
          /note="CpG island (%GC=75.0, o/e=0.98, #CpGs=206)"
          103525..103701
          /note="Sequence derived from PCR product of project DNA."
          103791..104120
          /note="Sequence derived from PCR product of project DNA."
          105003..105008
          /note="Sequence derived from one plasmid subclone."
          107451..107523
          /product="tRNA-Ser"
          /note="Likely pseudogene (HMM Sc=32.43 / Sec struct
            Sc=7.29)"
          113983..114055
          /product="tRNA-Ser"
          /note="Likely pseudogene (HMM Sc=40.24 / Sec struct
            Sc=7.11)"
          complement(118864..118936)
          /product="tRNA-Ser"
          /note="Likely pseudogene (HMM Sc=39.14 / Sec struct
            Sc=14.97)"

```

ORIGIN

```

Query Match      76.8%; Score 19.2; DB 9; Length 129045;
Best Local Similarity 87.5%; Pred. No. 3.7e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 1 AAAAAACATGACCACTGGGCTG 24
    ||||| ||||| ||||| |||||
Db 42068 AAAAAAAGTCCAACCTGGGCTG 42091

RESULT 46
AC135550 129968 bp DNA linear HTG 25-JAN-2003
LOCUS Pan troglodytes clone RP43-65H8, WORKING DRAFT SEQUENCE, 4 ordered
DEFINITION pieces.
AC135550
VERSION 1 (bases 1 to 129968)
KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT.
SOURCE Pan troglodytes (Chimpanzee)
ORGANISM Pan troglodytes
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Pan.
REFERENCE
  1 (bases 1 to 129968)
AUTHORS Akhter,N., Antonellis,A., Ayele,K., Beckstrom-Sternberg,S.M.,
  Benjamin,B., Blakesley,R.W., Bouffard,G.G., Brinkley,C., Brooks,S.,
  Carlaga,K., Coleman,B., Engle,J., Granite,S., Guan,X., Gupta,J.,
  Haghighi,P., Han,J., Hansen,N., Ho,S.-L., Idol,J.R., Karlins,E.,
  Laric,P., Lee-Lin,S.-Q., Legaspi,R., Maduro,Q.L., Maduro,V.B.,
  Margulies,E.H., Masiello,C., Maskeri,B., McDowell,J.,
  Paguirigan,C., Pearson,R., Portnoy,M.E., Prasad,A.,
  Reddix-Dugue,N., Schandler,K., Schueler,M.G., Sison,C.,
  Stantripoop,S., Thomas,J.W., Thomas,P.J., Touchman,J.W., Vogt,J.L.,
  Wetherby,K.D., Wiggins,L., Young,A. and Green,E.D.
  NISC Comparative Sequencing Initiative
  Unpublished
REFERENCE
  2 (bases 1 to 129968)
  Green,E.D.
  Direct Submission
  Submitted (18-OCT-2002) NIH Intramural Sequencing Center, 8717
  Grovemont Circle, Gaithersburg, MD 20877, USA
  3 (bases 1 to 129968)
  Green,E.D.
  Direct Submission
  Submitted (25-JAN-2003) NIH Intramural Sequencing Center, 8717
  Grovemont Circle, Gaithersburg, MD 20877, USA
  On Jan 25, 2003 this sequence version replaced gi:24110979.
  ----- Genome Center
  Center: NIH Intramural Sequencing Center
  Center code: NISC
  Web site: http://www.nisc.nih.gov
  Contact: nisc\_zoo@nhgri.nih.gov
  ----- Project Information
  Center project name: cmd
  Center clone name: 065H08

```

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

```

----- Summary Statistics
Sequencing vector: plasmid, n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 129301 bases at least Q40
Consensus quality: 129516 bases at least Q30
Consensus quality: 129643 bases at least Q20
Insert size: 129000; agarose-fp
Insert size: 129668; sum-of-contigs
Quality coverage: 10.58x in Q20 bases; agarose-fp

```

Quality coverage: 10.52x in Q20 bases; sum-of-contigs

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 4 contigs. Gaps between the contigs
 * are represented as runs of N. The order of the pieces
 * is believed to be correct as given, however the sizes
 * of the gaps between them are based on estimates that have
 * provided by the submitter.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.
 * 1 6330: contig of 6330 bp in length
 * 6331 6430: gap of unknown length
 * 6431 9836: contig of 3406 bp in length
 * 9837 9936: gap of unknown length
 * 9937 122751: contig of 112815 bp in length
 * 122752 122851: gap of unknown length
 * 122852 122968: contig of 7117 bp in length.

FEATURES

Location/Qualifiers
 1..122968
 /organism="pan troglodytes"
 /mol_type="genomic DNA"
 /db_xref="taxon:9598"
 /clone="RP43-65H8"
 /clone_lib="RP43"

source

misc_feature

1..6330
 /note="assembly_fragment
 clone_end:SP6
 vector_side:left"
 6331..6430

gap

/estimated_length=unknown
 6431..9836

misc_feature

/note="assembly_fragment"
 9837..9936

gap

/estimated_length=unknown
 9937..122751

misc_feature

/note="assembly_fragment"
 79726..122968

misc_feature

/note="clone overlaps with GenBank Accession Number
 AC117936 clone RP43-90G9 (center project name cmc)"
 122752..122851

gap

/estimated_length=unknown
 122852..122968

misc_feature

/note="assembly_fragment
 clone_end:17
 vector_side:right"

ORIGIN

Query Match 76.8%; Score 19.2; DB 14; Length 122968;
 Best Local Similarity 87.5%; Pred. No. 3.7e+02;
 Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAACATGACCACTGGGCTG 24

Db 36236 AAAAAACATGACCTACTGGGCTG 36259

RESULT 47

AC144765/c AC144765 130836 bp DNA linear PLN 14-JUN-2005
 LOCUS Medicago truncatula clone mth2-35011, complete sequence.
 DEFINITION AC144765

ACCESSION AC144765

VERSION AC144765.19 GI:67625917

KEYWORDS HTG.

SOURCE Medicago truncatula (barrel medic)

ORGANISM Medicago truncatula

Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
 Medicago.

REFERENCE 1 (bases 1 to 130836)

AUTHORS Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
 Cook,D., Kim,D. and Roe,B.A.

Medicago truncatula BAC Clone mth2-35011

Unpublished

2 (bases 1 to 130836)

Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
 Cook,D., Kim,D. and Roe,B.A.

Direct Submission

Submitted (16-MAY-2003) Department Of Chemistry And Biochemistry,
 The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
 OK 73019, USA

REFERENCE

AUTHORS

3 (bases 1 to 130836)
 Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
 Cook,D., Kim,D. and Roe,B.A.

Direct Submission

Submitted (06-MAY-2005) Department Of Chemistry And Biochemistry,
 The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
 OK 73019, USA

REFERENCE

AUTHORS

4 (bases 1 to 130836)
 Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
 Cook,D., Kim,D. and Roe,B.A.

Direct Submission

Submitted (07-MAY-2005) Department Of Chemistry And Biochemistry,
 The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
 OK 73019, USA

REFERENCE

AUTHORS

5 (bases 1 to 130836)
 Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
 Cook,D., Kim,D. and Roe,B.A.

Direct Submission

Submitted (10-MAY-2005) Department Of Chemistry And Biochemistry,
 The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
 OK 73019, USA

REFERENCE

AUTHORS

6 (bases 1 to 130836)
 Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
 Cook,D., Kim,D. and Roe,B.A.

Direct Submission

Submitted (14-MAY-2005) Department Of Chemistry And Biochemistry,
 The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
 OK 73019, USA

REFERENCE

AUTHORS

7 (bases 1 to 130836)
 Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
 Cook,D., Kim,D. and Roe,B.A.

Direct Submission

Submitted (14-JUN-2005) Department Of Chemistry And Biochemistry,
 The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
 OK 73019, USA

COMMENT

On Jun 14, 2005 this sequence version replaced gi:65133652.

----- Genome Center

Center: Department Of Chemistry And Biochemistry

The University Of Oklahoma

Center code:UOKNOR

FEATURES

source

Location/Qualifiers
 1..130836
 /organism="Medicago truncatula"
 /mol_type="genomic DNA"
 /db_xref="taxon:3880"
 /clone="mth2-35011"
 /clone_lib="Medicago truncatula BAC library H2"

ORIGIN

Query Match 76.8%; Score 19.2; DB 15; Length 130836;
 Best Local Similarity 87.5%; Pred. No. 3.7e+02;
 Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAACATGACCACTGGGCTG 24

Db 70159 AAAAAACATGACCAAGTGGCTG 70136

RESULT 48

AC145828

LOCUS

DEFINITION

ACCESSION

139712 bp DNA linear PRI 26-SEP-2003
 Pan troglodytes BAC clone RP43-7E4 from 7, complete sequence.
 AC145828

LOCUS AC027584 144003 bp DNA linear PRI 25-NOV-2001
 DEFINITION Homo sapiens chromosome 18, clone RP11-214L13, complete sequence.
 AC027584
 VERSION AC027584.5 GI:16596606
 HTG.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Hominidae; Homo.
 1 (bases 1 to 144003)
 Birren,B., Linton,L., Nusbaum,C. and Lander,E.
 Homo sapiens chromosome 18, clone RP11-214L13
 Unpublished
 2 (bases 1 to 144003)
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
 Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
 Boguslavskiy,L., Boukhgalter,B., Brown,A., Burkett,G.,
 Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
 Collymore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S.,
 Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
 Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
 Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
 Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
 Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczkv,J.,
 Levine,R., Liu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
 McCarthy,M., McEwan,P., McGuck,A., McKernan,K., McPheeters,R.,
 Meldrum,J., Meneus,L., Mihova,T., Miranda,C., Mienga,V., Morrow,J.,
 Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
 O'Neil,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
 Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
 Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
 Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
 Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
 Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
 Young,G., Zainoun,J., Zimmer,A. and Zody,M.
 Direct Submission
 Submitted (30-MAR-2000) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 3 (bases 1 to 144003)
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
 Anderson,S., Barna,N., Bastien,V., Boguslavskiy,L., Boukhgalter,B.,
 Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
 Brown,C., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
 Choepel,Y., Colangelo,M., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
 Cooke,P., Dearellano,K., Dewar,K., FitzHugh,W., Gage,D., Galagan,J.,
 Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
 Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
 Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K.,
 Lamazares,R., Landers,T., Lehoczkv,J., Levine,R., Liu,G.,
 Maclean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
 McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrum,J.,
 Meneus,L., Mihova,T., Mienga,V., Murphy,T., Naylor,J., Nguyen,C.,
 Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
 Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
 Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
 Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
 Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
 Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
 Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
 Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
 Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
 Direct Submission
 Submitted (25-NOV-2001) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Nov 2, 2001 this sequence version replaced gi:14994164.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html
 ----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: http://www-seq.wi.mit.edu

Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L9024
 Center clone name: 214_L_13

 FEATURES
 source
 1. 144003
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="18"
 /map="18"
 /clone="RP11-214L13"
 /clone_lib="RPCI-11 Human Male BAC"
 230..289
 /rpt_family="L2"
 complement(344..691)
 /rpt_family="L1MEC"
 complement(736..1232)
 /rpt_family="L1MEC"
 complement(1287..2755)
 /rpt_family="L1PA5"
 complement(2756..3066)
 /rpt_family="AluYa5"
 complement(3067..6236)
 /rpt_family="L1PA5"
 6239..6342
 /rpt_family="CT-rich"
 complement(6365..6657)
 /rpt_family="AluX"
 complement(6673..8150)
 /rpt_family="L1MEC"
 complement(9977..10170)
 /rpt_family="MIR"
 10843..10945
 /rpt_family="MIR3"
 11285..11322
 /rpt_family="AT-rich"
 complement(11536..11722)
 /rpt_family="L1MC4a"
 complement(11788..12154)
 /rpt_family="HAL1b"
 complement(12247..12933)
 /rpt_family="HAL1b"
 13591..13614
 /rpt_family="(TA)n"
 complement(13875..14038)
 /rpt_family="L1MC4a"
 14173..14230
 /rpt_family="MIR3"
 15625..15796
 /rpt_family="MERS7B"
 15828..15901
 /rpt_family="MERS7A"
 complement(15903..16193)
 /rpt_family="AluJb"
 17096..17153
 /rpt_family="MERS7A-int"
 complement(17154..17448)
 /rpt_family="AluSq"
 17449..17529
 /rpt_family="MERS7A-int"
 17690..18056
 /rpt_family="MERS7A-int"
 18197..18302
 /rpt_family="MER34"
 18303..18368
 /rpt_family="MER34"
 18588..19552
 /rpt_family="MER57-int"
 19535..20187
 /rpt_family="MER57-int"
 20586..20784
 repeat_region

repeat_region /rpt family="MER57-int"
21307..22250
repeat_region /rpt family="MER57-int"
2251..22605
repeat_region /rpt family="MER57A"
23090..23121
repeat_region /rpt family="(TA)n"
complement(23122..23420)
repeat_region /rpt family="AluJo"
23857..24010
repeat_region /rpt family="MIR3"
24147..24211
repeat_region /rpt family="MER31A"
24917..24940
repeat_region /rpt family="AT_rich"
25532..25715
repeat_region /rpt family="MER20"
26047..26085
repeat_region /rpt family="AT_rich"
complement(26320..26452)
repeat_region /rpt family="LINE"
26603..26630
repeat_region /rpt family="AT_rich"
complement(27226..27530)
repeat_region /rpt family="AluSx"
complement(28029..28085)
repeat_region /rpt family="LINE"
complement(28107..29600)
repeat_region /rpt family="LMB3"
complement(30124..30460)
repeat_region /rpt family="LMB3"
complement(31116..31622)
repeat_region /rpt family="LMB3"
32859..32895
repeat_region /rpt family="AT_rich"
33018..33201
repeat_region /rpt family="MER58A"
34924..35065
repeat_region /rpt family="MIR"
complement(35066..35294)
repeat_region /rpt family="MER3"
35295..35383
repeat_region /rpt family="MIR"
35556..35815
repeat_region /rpt family="MIR"
complement(36703..36851)
repeat_region /rpt family="L2"
37033..37062
repeat_region /rpt family="AT_rich"
37142..37749
repeat_region /rpt family="L1PA15"
37792..38085
repeat_region /rpt family="AluJo"
39160..39343

Query Match 76.8%; Score 19.2; DB 8; Length 144003;
Best Local Similarity 87.5%; Pred. No. 3.7e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAACAATGACCAACTGGGCTGT 25
|||||
Db 48295 AAAAGCANTGACCAACTGGGCTGT 48272

Search completed: February 3, 2006, 21:26:14
Job time : 877 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 3, 2006, 10:53:33 ; Search time 203.111 Seconds
(without alignments)
820.326 Million cell updates/sec

Title: US-10-719-900-5

Perfect score: 25

Sequence: 1 aaaaacaatgacccaactgggtgt 25

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database :

- N Geneseq_21:*
- 1: Geneseqn1980s.*
 - 2: Geneseqn1990s.*
 - 3: Geneseqn2000s.*
 - 4: Geneseqn2001as.*
 - 5: Geneseqn2001bs.*
 - 6: Geneseqn2002as.*
 - 7: Geneseqn2002bs.*
 - 8: Geneseqn2003as.*
 - 9: Geneseqn2003bs.*
 - 10: Geneseqn2003cs.*
 - 11: Geneseqn2003ds.*
 - 12: Geneseqn2004as.*
 - 13: Geneseqn2004bs.*
 - 14: Geneseqn2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20.8	83.2	57248	6	ABK83563 Human CDN
2	19.2	76.8	600	4	Aak11812 Human bra
3	19.2	76.8	175603	12	ADQ97554 Mouse ca
C 4	18.6	75.2	6596	4	ABL08072 Drosophil
5	18.6	74.4	290547	13	ABD32598 Human cal
C 6	18.4	73.6	308	6	ABL62734 Breast ca
C 7	18.4	73.6	2146	12	ADQ22079 Human so
8	18.2	72.8	650	13	ACN62534 Cotton d
9	18.2	72.8	1003	8	ACA49099 Prokaryot
10	18.2	72.8	1047	8	ACA51846 Prokaryot
11	18.2	72.8	2319	13	ADX09399 Plant fu
C 12	18.2	72.8	54810	13	ABD32759 Mouse ca
C 13	18.2	72.8	66479	6	ABQ80567 Mutant hu
C 14	18.2	72.8	66479	6	ABQ80566 Mutant hu
C 15	18.2	72.8	66479	6	ABQ80568 Mutant hu
C 16	18.2	72.8	66479	6	ABQ80565 Human IKB
17	18.2	72.8	110000	13	ABD32804_4 Continuation (5 o
C 18	18	72.0	1522	6	ABZ16013 Arabidops
C 19	18	72.0	1522	8	ADA69122 Arabidops

93	17	68.0	441	5	ABV19324	Abv19324 Human pro
94	17	68.0	484	10	ABV5553	Abv5553 Toxicity-
95	17	68.0	492	5	ABV49098	Abv49098 Human pro
96	17	68.0	500	11	ADM65753	Adm65753 NRY polym
97	17	68.0	501	6	ABV53676	Abv53676 Plantanto
98	17	68.0	519	13	ADX31946	Adx31946 Plant ful
99	17	68.0	519	13	ABV55948	Abv55948 Human pro
100	17	68.0	613	5	ABV55948	Abv55948 Human pro
101	17	68.0	700	4	AHL92872	Ahl92872 Human Inf
102	17	68.0	728	4	ABL20019	Ab120019 Drosophil
103	17	68.0	761	12	ADO62623	Ado62623 Transcrip
104	17	68.0	765	12	ADQ17998	Adq17998 Human sof
105	17	68.0	984	6	ABO69809	Abg69809 Listeria
106	17	68.0	984	6	ABO68114	Abg68114 Listeria
107	17	68.0	984	8	ACA36535	Ac36535 Prokaryot
108	17	68.0	996	6	ABO67967	Abg67967 Listeria
109	17	68.0	996	6	ABO69994	Abg69994 Listeria
110	17	68.0	1077	14	ADW81848	Adw81848 MAP3K9 ma
111	17	68.0	1082	5	ABV28252	Abv28252 Human pro
112	17	68.0	1082	5	ABV22438	Abv22438 Human pro
113	17	68.0	1176	8	ACA33726	Ac33726 Prokaryot
114	17	68.0	1380	12	ADI29372	Adi29372 Mouse MAR
115	17	68.0	1674	13	ADT19380	Adt19380 Plant cDN
116	17	68.0	2000	7	ADZ74598	Adz74598 Arabidops
117	17	68.0	2015	6	ABK63485	Abk63485 Rat seque
118	17	68.0	2015	10	ABD57938	Abd57938 Toxicity-
119	17	68.0	2015	11	ADW21828	Adw21828 Rat hepat
120	17	68.0	2023	11	ADW22482	Adw22482 Rat hepat
121	17	68.0	2026	10	ADB58195	Abd58195 Toxicity-
122	17	68.0	2026	10	ABD52699	Abd52699 Primary r
123	17	68.0	2026	10	ABD52699	Abd52699 Primary r
124	17	68.0	2397	6	ABX13908	Abx13908 cDNA enco
125	17	68.0	3144	13	ADR07476	Adr07476 Full leng
126	17	68.0	3260	11	ADM01957	Adm01957 Human cDN
127	17	68.0	3377	4	ABL20018	Ab120018 Drosophil
128	17	68.0	3818	4	ABL18574	Ab118574 Drosophil
129	17	68.0	5508	10	ADE85828	Ad85828 Human pai
130	17	68.0	5508	10	ADE85827	Ad85827 Human pai
131	17	68.0	5508	10	ADE85824	Ad85824 Human pai
132	17	68.0	5508	10	ADE85826	Ad85826 Human pai
133	17	68.0	5775	6	ABO70927	Abg70927 Listeria
134	17	68.0	8189	4	ABL01988	Ab101988 Drosophil
135	17	68.0	8497	14	AEA61764	Aea61764 Streptoco
136	17	68.0	9905	6	AK98324	Aak98324 Human pur
137	17	68.0	9985	4	ABL05688	Ab105688 Drosophil
138	17	68.0	24888	13	ABD33500	Abd33500 Human can
139	17	68.0	28854	8	ABX95685	Abx95685 Human gen
140	17	68.0	33780	4	AAH24652	Aah24652 Nucleotid
141	17	68.0	40491	11	ACN44798	Acn44798 Human gen
142	17	68.0	42324	4	AAK84724	Aak84724 Human inm
143	17	68.0	104062	11	ACN44458	Acn44458 Human gen
144	17	68.0	110000	6	ABA03041_00	Ab303041 Listeria
145	17	68.0	110000	6	ABA03041_14	Ab303041 Listeria
146	17	68.0	117750	13	ABD32653	Abd32653 Human can
147	17	68.0	165554	11	ACN44108	Acn44108 Mouse gen
148	17	68.0	175378	14	ADV09445	Adv09445 Human ORP
149	17	68.0	256493	11	ACN44514	Acn44514 Human gen
150	16.8	67.2	885	4	AAH24195	Aah24195 Mouse age

ALIGNMENTS

RESULT 1
ABK83563
ID ABK83563 standard; cDNA; 57248 BP.
XX
AC ABK83563;

14-AUG-2002 (first entry)

DE Human cDNA differentially expressed in granulocytic cells #134.

XX

KW Human; ss; granulocytic cell; DNA chip; bacterial infection;
KW viral infection; parasitic infection; protozoal infection;
KW fungal infection; sterile inflammatory disease; psoriasis;
KW rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;
KW cardiac reperfusion injury; renal reperfusion injury; ARDS;
KW adult respiratory distress syndrome; inflammatory bowel disease;
KW Crohn's disease; ulcerative colitis; periodontal disease;
KW granulocyte activation; chronic inflammation; allergy.
OS Homo sapiens.
XX WO200228999-A2.
PN 11-APR-2002.
XX 03-OCT-2001; 2001WO-US030821.
XX 03-OCT-2000; 2000US-0237189P.
XX (GENE-) GENE LOGIC INC.
XX Beazer-Barclay Y, Weissman SM, Yamaga S, Vockley J;
XX WPI; 2002-435328/46.
XX Detecting granulocyte activation by detecting differential expression of
XX genes associated with granulocyte activation, which serves as diagnostic
XX markers that is useful for monitoring disease states and drug toxicity.
XX Claim 1; SEQ ID NO 134; 114pp; English.
XX The invention relates to detecting (M1) granulocyte (GC) activation
XX (GCA), by detecting the level of expression of gene(s) (Gs) identified by
XX DNA chip analysis as given in the specification, and comparing the
XX expression level to an expression level in an unactivated GC, where
XX differential expression of Gs is indicative of GCA. Also included are
XX modulating (M2) GCA by contacting GC with an agent that alters the
XX expression of at least one gene in Gs; (2) screening (M3) for an agent
XX capable of modulating GCA or an inflammation (especially chronic) in a
XX tissue, an allergic response in a subject, exposure of a subject to a
XX pathogen or sterile inflammatory disease using the gene expression
XX profile; (3) detecting (M4) an inflammation (especially chronic) in a
XX tissue, an allergic response in a subject, exposure of a subject to a
XX pathogen or sterile inflammatory disease, by detecting the level of
XX expression in a sample of the tissue of gene(s) from Gs, where the level
XX of expression of the gene is indicative of inflammation; (4) treating
XX (M5) an inflammation (especially chronic) or in a tissue, an allergic
XX response in a subject, exposure of a subject to a pathogen or sterile
XX inflammatory disease, by contacting a tissue having inflammation with an
XX agent that modulates the expression of gene(s) from Gs in the tissue. M1
XX is useful for detecting GCA; M2 is useful for modulating GCA; M3 is useful
XX for screening an agent capable of modulating GCA preferably in an
XX inflammation in a tissue; M4 is useful for detecting an inflammation
XX (especially chronic) in a tissue, an allergic response in a subject,
XX exposure of a subject to a pathogen or sterile inflammatory disease (e.g.
XX psoriasis, rheumatoid arthritis, glomerulonephritis, asthma, thrombosis,
XX cardiac reperfusion injury, renal reperfusion injury, ARDS, adult
XX respiratory distress syndrome, inflammatory bowel disease, Crohn's
XX disease, ulcerative colitis, periodontal disease; also bacterial
XX infection, viral infection, parasitic infection, protozoal infection,
XX fungal infection and M5 is useful for treating one of the above
XX conditions. The present sequence represents a gene differentially
XX expressed in granulocytes. Note: The sequence data for this patent did
XX not form part of the printed specification, but was obtained in
XX electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 57248 BP; 15003 A; 13601 C; 13307 G; 15337 T; 0 U; 0 Other;

Query Match 83.2%; Score 20.8; DB 6; Length 57248;
Best Local Similarity 91.7%; Pred. No. 37;
Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
QY 1 AAAAAACAATGACCAACTGGGCTG 24
Db 42908 AAAAAACAATGTCCTGCTG 42931
|||||

RESULT 2
AAK11812
ID AAK11812 standard; DNA; 600 BP.
XX
AC AAK11812;
XX
DT 05-NOV-2001 (first entry)
XX
DE Human brain expressed single exon probe SEQ ID NO: 11803.
XX
KW Human; brain expressed exon; gene expression analysis; probe; microarray;
KW Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer;
KW ss.
XX
OS Homo sapiens.
XX
XX
XX WO200157275-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US000667.
XX
XX 04-FEB-2000; 2000US-0180312P.
XX
XX 26-MAY-2000; 2000US-0207456P.
XX
XX 30-JUN-2000; 2000US-00608408.
XX
XX 03-AUG-2000; 2000US-00632366.
XX
XX 21-SEP-2000; 2000US-0234687P.
XX
XX 27-SEP-2000; 2000US-0236359P.
XX
XX 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-483446/52.
XX
XX Single exon nucleic acid probes for analyzing gene expression in human
XX brains.
XX
XX Example 4; SEQ ID NO 11803; 650pp + Sequence Listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
XX probes which are derived from genomic sequences expressed in the human
XX brain. They can be used to measure gene expression in brain cell samples,
XX which may enable the diagnosis and improved treatment of nervous system
XX diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
XX epilepsy and cancers. The present sequence is one of the probes of the
XX invention.
XX
XX Sequence 600 BP; 187 A; 104 C; 123 G; 186 T; 0 U; 0 Other;
XX
Query Match 76.8%; Score 19.2; DB 4; Length 600;
Best Local Similarity 87.5%; Pred. No. 98;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAAACAATGACCAACTGGGCTG 25
Db 370 AAAAAACAATGACCAACTGGTTGT 393
|||||

RESULT 3
ADQ97554
ID ADQ97554 standard; DNA; 175603 BP.
XX
AC ADQ97554;
XX
XX
XX 07-OCT-2004 (first entry)
XX
XX
```

```
DE Mouse cancer associated sequence MD09-013, SEQ ID 531.
XX
KW Cytostatic; Gene Therapy; cancer; leukemia; lymphoma; Mouse; ds.
XX
OS Mus musculus.
XX
XX WO2004060304-A2.
XX
XX PD 22-JUL-2004.
XX
XX PF 22-DEC-2003; 2003WO-US041389.
XX
XX PR 27-DEC-2002; 2002US-00330773.
XX
XX PA (SAGR-) SAGRES DISCOVERY INC.
XX
XX PI Morris DW, Malandro MS;
XX
XX WPI; 2004-543781/52.
XX
XX New isolated cancer associated nucleic acids comprising at least 10
XX contiguous nucleotides, useful for diagnosing, preventing and/or treating
XX cancers such as leukemia and lymphoma.
XX
XX Claim 1; SEQ ID NO 531; 199pp; English.
XX
XX The present invention relates to cancer associated sequences (ADQ97025-
XX ADQ98004). The sequences are useful for the diagnosis, prevention and/or
XX treatment of cancer, such as leukemia and lymphoma. Note: The sequence
XX data for this patent did not form part of the printed specification, but
XX was obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 175603 BP; 48715 A; 34685 C; 35401 G; 51161 T; 0 U; 5641 Other;
XX
Query Match 76.8%; Score 19.2; DB 12; Length 175603;
Best Local Similarity 87.5%; Pred. No. 2.4e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAAACAATGACCAACTGGGCTGT 25
Db 147030 AAAAAACAATGACCAAGCGGTTGT 147053
|||||

RESULT 4
ABL08072/c
ID ABL08072 standard; cDNA; 6596 BP.
XX
AC ABL08072;
XX
XX DT 26-MAR-2002 (first entry)
XX
XX DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 18698.
XX
XX Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical; gene; ss.
XX
XX Drosophila melanogaster.
XX
XX WO200171042-A2.
XX
XX PD 27-SEP-2001.
XX
XX PF 23-MAR-2001; 2001WO-US009231.
XX
XX PR 23-MAR-2000; 2000US-0191637P.
XX
XX PR 11-JUL-2000; 2000US-00614150.
XX
XX PA (PEKE ) PE CORP NY.
XX
XX PI Venter JC, Adams M, Li PWD, Myers EW;
XX
XX WPI; 2001-656860/75.
XX
```

```
DR P-PSDB; ABB63969.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX
PS Claim 1; SEQ ID NO 18698; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 6596 BP; 1724 A; 1400 C; 1614 G; 1858 T; 0 U; 0 Other;
Query Match 75.2%; Score 18.8; DB 4; Length 6596;
Best Local Similarity 90.9%; Pred. No. 2.2e+02;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 2 AAAAAAATGACCAACTGGCT 23
DB 6184 AAAAAAATGACCAACTGGACT 6163
RESULT 5
ABD32598
ID ABD32598 standard; DNA; 290547 BP.
XX
AC ABD32598;
XX
XX 18-NOV-2004 (first entry)
DE Human cancer-associated genomic DNA HD7-220.
KW Human; ds; cancer-associated protein; gene; cytostatic; cancer;
KW leukaemia; lymphoma; CAP.
XX Homo sapiens.
XX WO2004074320-A2.
XX 02-SEP-2004.
XX 17-FEB-2004; 2004WO-US004730.
XX 14-FEB-2003; 2003US-00367094.
XX 14-MAR-2003; 2003US-00388838.
XX 15-APR-2003; 2003US-00417375.
XX 13-JUN-2003; 2003US-00461862.
XX 15-SEP-2003; 2003US-00663431.
XX 15-DEC-2003; 2003US-00737318.
XX (SAGR-) SAGRES DISCOVERY INC.
XX
XX Morris DW, Morris DW, Malandro MS;
XX WPI; 2004-652914/63.
XX
XX New isolated cancer-associated polynucleotides and polypeptides useful
XX for diagnosing, preventing or treating cancers, especially lymphoma and
XX leukemia, or in screening for agents that modulate cancer.
XX claim 16; seqid 103; 310pp; English.
XX
XX The invention relates to an isolated nucleic acid comprising at least 10
XX contiguous nucleotides of any of the 233 polynucleotide sequences given
XX in the specification, or its complement. The nucleic acids encode cancer-
```

```
CC associated proteins. Also included are an expression vector comprising
CC the isolated nucleic acid cited above, a host cell comprising the above
CC a recombinant nucleic acid or expression vector, a microarray for detecting
CC a cancer-associated (CA) nucleic acid comprising at least one probe
CC comprising at least 10 contiguous nucleotides of any of the above-
CC mentioned nucleotide sequences, an isolated polypeptide (encoded within
CC an open reading frame of a CA sequence selected from any of the 95
CC polynucleotide sequences as mentioned in the specification, or its
CC complement), an isolated antibody, (or its antigen binding fragment) that
CC binds to the above polypeptide, a hybridoma that produces the above
CC monoclonal antibody, a pharmaceutical composition comprising the above
CC antibody and a pharmaceutical excipient, a kit for detecting cancer
CC cells (comprising the antibody cited above, methods for diagnosing cancer
CC or for detecting the presence or absence of cancer cells in an
CC individual, a method for inhibiting growth of cancer cells in an
CC individual, a method for delivering a therapeutic agent to cancer cells
CC in an individual, an electronic library comprising the above
CC polynucleotide or polypeptide (or their fragments), methods of screening
CC for anticancer activity or for a bioactive agent capable of modulating
CC the activity of a CA protein (CAP), methods for detecting cancer,
CC associated with expression of a polypeptide in a test cell sample, a
CC method for treating cancers and a method for inhibiting the expression of
CC CA gene in a cell. The composition and methods are useful for detecting,
CC diagnosing, preventing and treating cancers, especially lymphoma and
CC leukaemia. These may also be used in screening for agents that modulate
CC cancer. The present sequence is a human CAP genomic sequence. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 290547 BP; 78477 A; 46715 C; 46576 G; 86851 T; 0 U; 31928 Other;
Query Match 74.4%; Score 18.6; DB 13; Length 290547;
Best Local Similarity 84.0%; Pred. No. 4.8e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 AAAAAACAATGACCAACTGGCTGT 25
DB 26055 AAAAAAATATCAACAACCTGGGCTGT 26079
RESULT 6
ABL62734/C
ID ABL62734 standard; DNA; 308 BP.
XX
AC ABL62734;
XX
XX 15-MAY-2002 (first entry)
XX Breast cancer related gene sequence SEQ ID NO:1071.
XX
XX Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
XX stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
XX cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
XX gene; ds.
XX Homo sapiens.
XX WO200194629-A2.
XX 13-DEC-2001.
XX 30-MAY-2001; 2001WO-US010838.
XX 05-JUN-2000; 2000US-0209473P.
XX 05-JUN-2000; 2000US-0209531P.
XX 18-SEP-2000; 2000US-0233133P.
XX 18-SEP-2000; 2000US-0233617P.
XX 20-SEP-2000; 2000US-0234009P.
XX 20-SEP-2000; 2000US-0234034P.
XX 20-SEP-2000; 2000US-0234052P.
XX 22-SEP-2000; 2000US-0234509P.
XX 22-SEP-2000; 2000US-0234567P.
```

25-SEP-2000; 2000US-0234923P.
25-SEP-2000; 2000US-0234924P.
25-SEP-2000; 2000US-0235077P.
25-SEP-2000; 2000US-0235082P.
25-SEP-2000; 2000US-0235134P.
25-SEP-2000; 2000US-0235280P.
26-SEP-2000; 2000US-0235637P.
26-SEP-2000; 2000US-0235638P.
27-SEP-2000; 2000US-0235711P.
27-SEP-2000; 2000US-0235720P.
27-SEP-2000; 2000US-0235840P.
27-SEP-2000; 2000US-0235863P.
28-SEP-2000; 2000US-0236028P.
28-SEP-2000; 2000US-0236032P.
28-SEP-2000; 2000US-0236033P.
28-SEP-2000; 2000US-0236034P.
28-SEP-2000; 2000US-0236109P.
28-SEP-2000; 2000US-0236111P.
29-SEP-2000; 2000US-0236842P.
29-SEP-2000; 2000US-0236891P.
02-OCT-2000; 2000US-0237172P.
02-OCT-2000; 2000US-0237173P.
02-OCT-2000; 2000US-0237278P.
02-OCT-2000; 2000US-0237294P.
02-OCT-2000; 2000US-0237295P.
02-OCT-2000; 2000US-0237316P.
03-OCT-2000; 2000US-0237425P.
03-OCT-2000; 2000US-0237598P.
03-OCT-2000; 2000US-0237604P.
03-OCT-2000; 2000US-0237606P.
03-OCT-2000; 2000US-0237608P.
01-NOV-2000; 2000US-0244867P.
01-NOV-2000; 2000US-0245084P.
XX
PA (AVAL-) AVALON PHARM.
XX
PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
XX
PI WPI; 2002-188264/24.
XX
PT Screening for anti-neoplastic agent involves exposing cells to a chemical
PT agent to be tested for anti-neoplastic activity, and determining a change
PT in expression of a gene of a signature gene set.
XX
PS Claim 1; SEQ ID NO 1071; 44pp; English.
XX
CC The present invention describes a method (M1) for screening for an anti-
CC neoplastic agent. The method involves exposing cells to a chemical agent
CC to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening an
CC anti-neoplastic agent, and can be used for producing a product which is
CC the data collected with respect to the anti-neoplastic agent as a result
CC of M1, and the data is sufficient to convey the chemical structure and/or
CC properties of the agent. M1 can be used in the treatment of cancer such
CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,
CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilms'
CC tumour
XX
SQ Sequence 308 BP; 71 A; 64 C; 67 G; 106 T; 0 U; 0 Other;
Query Match 73.6%; Score 18.4; DB 6; Length 308;
Best Local Similarity 95.0%; Pred. No. 2.1e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 AAAAAACAATGACCAACTGG 20
|||||

21 AAAAAACAATGTCCAACTGG 2
Db
RESULT 7
ADQ22079/c
ID ADQ22079 standard; DNA; 2146 BP.
XX
AC ADQ22079;
XX
DT 26-AUG-2004 (first entry)
XX
DE Human soft tissue sarcoma-upregulated DNA - SEQ ID 4899.
XX
KW soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human;
KW ds.
XX
OS Homo sapiens.
XX
FN WO2004048938-A2.
XX
PD 10-JUN-2004.
XX
PF 26-NOV-2003; 2003WO-US038193.
XX
PR 26-NOV-2002; 2002US-0429739P.
XX
PA (PROT-) PROTEIN DESIGN LABS INC.
XX
PI Aziz N, Ginsburg WM, Zlotnik A;
XX
XX WPI; 2004-441208/41.
XX
PT Early detection of soft tissue sarcoma comprises determining expression
PT of a gene in a first soft tissue sample and a normal soft tissue sample
PT and comparing the gene expression, also useful in treating soft tissue
PT sarcoma.
XX
PS Example 2; SEQ ID NO 4899; 210pp; English.
XX
CC The invention relates to a novel method for detecting soft tissue sarcoma
CC which comprises obtaining a first soft tissue sample from an individual
CC and a normal soft tissue sample from the same or different individual,
CC determining the expression of a gene in both samples and comparing the
CC expression of the gene in both soft tissue samples, where a higher level
CC of protein expression in the first soft tissue sample indicates the
CC presence of soft tissue sarcoma. The method of the invention has
CC cytostatic applications and may be useful for detecting soft tissue
CC sarcoma, possibly via gene therapy or vaccine production. The nucleic
CC acid sequences may be useful in diagnostic and screening applications.
CC The current sequence is that of a human soft tissue sarcoma-upregulated
CC DNA of the invention. The current sequence is not shown within the
CC specification per se but was submitted in CD format by the inventor.
XX
SQ Sequence 2146 BP; 536 A; 443 C; 428 G; 739 T; 0 U; 0 Other;
Query Match 73.6%; Score 18.4; DB 12; Length 2146;
Best Local Similarity 95.0%; Pred. No. 2.8e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 AAAAAACAATGACCAACTGG 20
|||||
Db 679 AAAAAACAATGTCCAACTGG 660
RESULT 8
ACN62534
ID ACN62534 standard; cDNA; 650 BP.
XX
AC ACN62534;
XX
DT 02-DEC-2004 (first entry)
XX
DE Cotton developing fibre EST Clone ID: L1B3830-001-Q1-K6-E8, SEQ:17315.

XX Cotton; plant; EST; expressed sequence tag; transgenic plant;
KW developing fibre; variety Nucleon338; library LIB3830; molecular tag;
KW molecular marker; genetic mapping; molecular mapping; seed germination;
KW plant growth; plant quality; plant yield; plant breeding;
KW tissue printing; ss.
XX
OS Gossypium hirsutum.
XX
XX US2004123340-A1.
XX
XX 24-JUN-2004.
XX
XX 12-DEC-2001; 2001US-00021323.
XX
XX 14-DEC-2000; 2000US-0255619P.
XX
XX (DEIK/) DEIKMAN J.
PA (FENG/) FENG P C C.
PA (FINC/) FINCHER K L.
PA (ZIEG/) ZIEGLER T E.
XX
XX Deikman J, Feng PCC, Fincher KL, Ziegler TE;
PI
XX
XX WPI; 2004-479808/45.
XX
XX New isolated nucleic acid molecule that encodes a plant protein or its
PT fragment, useful for isolating a variety of agronomically significant
PT genes associated with plant growth, quality or yield, and as molecular
PT tags to map genes.
XX
XX Claim 1; SEQ ID NO 17315; 34pp; English.
XX
XX The invention relates to 17880 cotton expressed sequence tags (ESTs;
CC ACN45220-ACN63099). The ESTs were isolated from cDNA libraries generated
CC from primed or non-primed seeds from variety DP50B, mature seeds from
CC variety Coker 312 Boswell 96 Field, and androecium tissue, gynoceum
CC tissue, developing fibres, carpel walls and septa from variety
CC Nucleon338. The invention also relates to substantially purified
CC proteins or their fragments encoded by nucleic acid molecules of the
CC invention, and to transformed plants having a nucleic acid construct
CC comprising a nucleic acid of the invention. The cotton ESTs are useful as
CC molecular tags to isolate genetic regions, to isolate genes, to map
CC genes, to determine gene function, and to determine whether genes are
CC members of a particular gene family. The nucleic acid molecules may be
CC used for isolating a variety of agronomically significant genes
CC associated with plant growth, quality, yield, and could also serve as
CC links in metabolic and catabolic pathways. The nucleic acid molecules are
CC also useful for identifying genes important in initiating and maintaining
CC seed germination or that may be used to mitigate stresses encountered
CC during seed germination. The ESTs additionally enable the acquisition of
CC promoters and cis-regulatory elements which will be useful to express
CC agronomically significant genes in these tissues and/or other tissues,
CC and also permits the acquisition of molecular markers useful in breeding
CC schemes, genetic and molecular mapping, and in cloning of agronomically
CC significant genes. The nucleic acid molecules are further useful for
CC detecting the expression level or pattern of a protein or mRNA and for
CC detecting the presence or quantity of a protein by tissue printing. The
CC present sequence represents a specifically claimed EST isolated from a
CC cotton variety Nucleon338 developing fibre cDNA library (LIB3830). The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the US
CC patent office at seqdata.uspto.gov/sequence.html?DocID=US20040123340
XX
SQ Sequence 650 BP; 193 A; 119 C; 136 G; 201 T; 0 U; 1 Other;
Query Match 72.8%; Score 18.2; DB 13; Length 650;
Best Local Similarity 87.0%; Pred. No. 2.9e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
OY 2 AAAACAATGACCAACTGGGCTG 24
Db 18 AAAACAATGACGACAGAGCTG 40

RESULT 9

ACA49099

ID ACA49099 standard; DNA; 1003 BP.

XX

AC ACA49099;

XX

DT 19-JUN-2003 (first entry)

XX

DE Prokaryotic essential gene #30756.

XX

KW Antisense; ds; prokaryotic essential gene; cell proliferation;

XX

KW drug design; gene.

XX

OS Salmonella paratyphi.

XX

PN WO200277183-A2.

XX

PD 03-OCT-2002.

XX

PF 21-MAR-2002; 2002WO-US0009107.

XX

PR 21-MAR-2001; 2001US-00815242.

XX

PR 06-SEP-2001; 2001US-00948993.

XX

PR 25-OCT-2001; 2001US-0342923P.

XX

PR 08-FEB-2002; 2002US-00072851.

XX

PR 06-MAR-2002; 2002US-0362699P.

XX

PA (ELIT-) ELITRA PHARM INC.

XX

PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;

XX

PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX

WPI; 2003-029926/02.

XX

P-PSDB; ABU45229.

XX

PT New antisense nucleic acids, useful for identifying proteins or screening

XX

PT for homologous nucleic acids, required for cellular proliferation to

XX

PT isolate candidate molecules for rational drug discovery programs.

XX

PS Claim 14; SEQ ID NO 36969; 1766pp; English.

XX

CC The invention relates to an isolated nucleic acid comprising any one of

XX

CC the 6213 antisense sequences given in the specification where expression

XX

CC of the nucleic acid inhibits proliferation of a cell. Also included are:

XX

CC (1) a vector comprising a promoter operably linked to the nucleic acid

XX

CC encoding a polypeptide whose expression is inhibited by the antisense

XX

CC nucleic acid; (2) a host cell containing the vector; (3) an isolated

XX

CC polypeptide or its fragment whose expression is inhibited by the

XX

CC antisense nucleic acid; (4) an antibody capable of specifically binding

XX

CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular

XX

CC proliferation or the activity of a gene in an operon required for

XX

CC proliferation; (7) identifying a compound that influences the activity of

XX

CC the gene product or that has an activity against a biological pathway;

XX

CC required for proliferation, or that inhibits cellular proliferation; (8)

XX

CC identifying a gene required for cellular proliferation or the biological

XX

CC pathway in which a proliferation-required gene or its gene product lies

XX

CC or a gene on which the test compound that inhibits proliferation of an

XX

CC organism acts; (9) manufacturing an antibiotic; (10) profiling a

XX

CC compound's activity; (11) a culture comprising strains in which the gene

XX

CC product is overexpressed or underexpressed; (12) determining the extent

XX

CC to which each of the strains is present in a culture or collection of

XX

CC strains; or (13) identifying the target of a compound that inhibits the

XX

CC proliferation of an organism. The antisense nucleic acids are useful for

XX

CC identifying proteins or screening for homologous nucleic acids required

XX

CC for cellular proliferation to isolate candidate molecules for rational

XX

CC drug discovery programs, or for screening homologous nucleic acids

XX

CC required for proliferation in cells other than S. aureus, S. typhimurium,

XX

CC K. pneumoniae or P. aeruginosa. The present sequence is one of the target

XX

CC prokaryotic essential genes. Note: The sequence data for this patent did

XX

CC not form part of the printed specification, but was obtained in

XX

CC electronic format directly from WIPO at

XX

CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1003 BP; 280 A; 249 C; 264 G; 210 T; 0 U; 0 Other;
Query Match 72.8%; Score 18.2; DB 8; Length 1003;
Best Local Similarity 87.0%; Pred. No. 3.1e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 2 AAAAACAATGACCAACTGGGCTG 24
||||| ||||| ||||| ||||| |||||
Db 399 AAAAACAATGACCAACTGGGCTG 421
RESULT 10
ACA51846
ID ACA51846 standard; DNA; 1047 BP.
XX
AC ACA51846;
XX
DT 19-JUN-2003 (first entry)
XX
DE Prokaryotic essential gene #33503.
XX
KW Antisense; ds; prokaryotic essential gene; cell proliferation;
KW drug design; gene.
XX
OS Salmomella typhi.
XX
PN WO200277183-A2.
XX
PD 03-OCT-2002.
XX
PP 21-MAR-2002; 2002WO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362899P.
XX
PA (BLIT-) ELITRA PHARM INC.
XX
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
XX WPI; 2003-029926/02.
DR P-PSDB; ABU47976.
XX
XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
XX Claim 14; SEQ ID NO 39716; 1766pp; English.
XX
XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
to which each of the strains is present in a culture or collection of
strains; or (13) identifying the target of a compound that inhibits the
proliferation of an organism. The antisense nucleic acids are useful for
identifying proteins or screening for homologous nucleic acids required
for cellular proliferation to isolate candidate molecules for rational
drug discovery programs, or for screening homologous nucleic acids
required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
K. pneumoniae or *P. aeruginosa*. The present sequence is one of the target
prokaryotic essential genes. Note: The sequence data for this patent did
not form part of the printed specification, but was obtained in an
electronic format directly from WIPO at
ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1047 BP; 295 A; 261 C; 273 G; 218 T; 0 U; 0 Other;
Query Match 72.8%; Score 18.2; DB 8; Length 1047;
Best Local Similarity 87.0%; Pred. No. 3.1e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 2 AAAAACAATGACCAACTGGGCTG 24
||||| ||||| ||||| ||||| |||||
Db 435 AAAAACAATGACCAACTGGGCTG 457
RESULT 11
ADX09399
ID ADX09399 standard; cDNA; 2319 BP.
XX
AC ADX09399;
XX
DT 21-APR-2005 (first entry)
XX
DE Plant full length insert polynucleotide seqid 3974.
XX
KW plant protectant; plant growth regulant; gene therapy; plant;
KW recombinant DNA construct; physical array; plant breeding marker;
KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
KW extreme osmotic condition; pathogen tolerance; pest tolerance;
KW growth rate; cell cycle pathway; disease resistance;
KW galactonnanan production; lignin production; plant growth regulator;
KW yield; plant growth; plant development; seed oil; protein yield;
KW protein content; gene; ss.
XX
XX Unidentified.
XX
XX US2004034888-A1.
XX
XX 19-FEB-2004.
XX
XX 28-APR-2003; 2003US-00425114.
XX
XX 06-MAY-1999; 99US-00304517.
XX
XX 05-NOV-2001; 2001US-00985678.
XX
XX (LIU/J) LIU J.
XX (ZHOU/J) ZHOU Y.
XX (KOVA/J) KOVALIC D K.
XX (SCRE/J) SCREEN S E.
XX (TABAS/J) TABASKA J E.
XX (CAO/Y) CAO Y.
XX
XX Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
XX WPI; 2004-180133/17.
XX
XX New recombinant DNA construct, useful for improving plant tolerance to
PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
PT pests, for conferring increased resistance to plant disease, or for
PT improving yield.
XX
XX Claim 1; SEQ ID NO 3974; 15pp; English.
XX
XX The invention describes a recombinant DNA construct comprising a

CC polynucleotide consisting of a sequence encoding an amino acid sequence
CC available in electronic form from the US patent office at
CC ftp.segdata.uspto.gov/sequence.html?docID:2004034888. The polynucleotide
CC of the invention are also useful in physical arrays of molecules and as
CC plant breeding markers. The recombinant DNA construct is useful for
CC improving plant tolerance to cold, heat, drought, herbicides, extreme
CC osmotic conditions, pathogens or pests, for manipulating growth rate in
CC plant cells by modification of the cell cycle pathway, for conferring
CC increased resistance to plant disease, for producing galactomanan,
CC lignin or plant growth regulators, for increasing the rate of homologous
CC recombination in plants, for improving yield by modification of
CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
CC or by providing improved plant growth and development under at least one
CC stress condition or for modifying seed oil or protein yield and/or
CC content. This sequence represents a plant full length insert
CC polynucleotide that can be used in the recombinant DNA construct of the
CC invention.
XX
SQ Sequence 2319 BP; 616 A; 487 C; 602 G; 614 T; 0 U; 0 Other;

Query Match 72.8%; Score 18.2; DB 13; Length 2319;
Best Local Similarity 87.0%; Pred. No. 3.5e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAACAATGACCAACTGGGCT 23
|||||
Db 2012 AAAAAACAATGACCAACTGGGCT 2034

RESULT 12
ABD32759/c
ID ABD32759 standard; DNA; 54810 BP.
XX
AC ABD32759;
XX
DT 18-NOV-2004 (first entry)
XX
DE Mouse cancer-associated genomic DNA MD16-030.
XX
KW Mouse; ds; cancer-associated protein; gene; cytostatic; cancer;
KW leukaemia; lymphoma; CAP.
XX
OS Mus musculus.
XX
PN WO2004074320-A2.
XX
PD 02-SEP-2004.
XX
PF 17-FEB-2004; 2004WO-US004730.
XX
PR 14-FEB-2003; 2003US-00367094.
PR 14-MAR-2003; 2003US-00388838.
PR 15-APR-2003; 2003US-00417375.
PR 13-JUN-2003; 2003US-00461862.
PR 15-SEP-2003; 2003US-00663431.
PR 15-DEC-2003; 2003US-00737318.
XX
PA (SAGR-) SAGRES DISCOVERY INC.
XX
PI Morris DW, Morris DW, Malandro MS;
XX
XX WPI; 2004-652914/63.
XX
XX New isolated cancer-associated polynucleotides and polypeptides useful
PT for diagnosing, preventing or treating cancers, especially lymphoma and
PT leukemia, or in screening for agents that modulate cancer.
XX
XX disclosure; seqid 365; 310pp; English.
XX
XX The invention relates to an isolated nucleic acid comprising at least 10
CC contiguous nucleotides of any of the 233 polynucleotide sequences given
CC in the specification, or its complement. The nucleic acids encode cancer-
CC associated proteins. Also included are an expression vector comprising

CC the isolated nucleic acid cited above, a host cell comprising the above
CC recombinant nucleic acid or expression vector, a microarray for detecting
CC a cancer-associated (CA) nucleic acid comprising at least one probe
CC comprising at least 10 contiguous nucleotides of any of the above-
CC mentioned nucleotide sequences, an isolated polypeptide (encoded within
CC an open reading frame of a CA sequence selected from any of the 95
CC polynucleotide sequences as mentioned in the specification, or its
CC complement), an isolated antibody, (or its antigen binding fragment) that
CC binds to the above polypeptide, a hybridoma that produces the above
CC monoclonal antibody, a pharmaceutical composition comprising the above
CC antibody and a pharmaceutical excipient, a kit for detecting cancer
CC cells (comprising the antibody cited above, methods for diagnosing cancer
CC or for detecting the presence or absence of cancer cells in an
CC individual, a method for inhibiting growth of cancer cells in an
CC individual, a method for delivering a therapeutic agent to cancer cells
CC in an individual, an electronic library comprising the above
CC polynucleotide or polypeptide (or their fragments), methods of screening
CC for anticancer activity or for a bioactive agent capable of modulating
CC the activity of a CA protein (CAP), methods for detecting cancer
CC associated with expression of a polypeptide in a test cell sample, a
CC method for treating cancers and a method for inhibiting the expression of
CC CA gene in a cell. The composition and methods are useful for detecting,
CC diagnosing, preventing and treating cancers, especially lymphoma and
CC leukaemia. These may also be used in screening for agents that modulate
CC cancer. The present sequence is a mouse CAP genomic sequence. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 54810 BP; 14231 A; 10582 C; 11156 G; 18235 T; 0 U; 606 Other;

Query Match 72.8%; Score 18.2; DB 13; Length 54810;
Best Local Similarity 87.0%; Pred. No. 5.7e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 AAAAAACAATGACCAACTGGGCTGT 25
|||||
Db 42877 ATAAACAATCAACCAACTGGGATGT 42855

RESULT 13
ABQ80567/c
ID ABQ80567 standard; DNA; 66479 BP.
XX
AC ABQ80567;
XX
DT 08-NOV-2002 (first entry)
XX
DE Mutant human IKBKAP gene #2.
XX
KW Human; IKBKAP; Familial Dysautonomia; FD; Riley-Day syndrome;
KW Hereditary Sensory and Autonomic Neuropathy Type III; carrier screening;
KW PD2; mutation; gene; chromosome 9q31; ds.
XX
OS Homo sapiens.
OS Synthetic.
XX
XX Key Location/Qualifiers
FH mutation replace(33714,G)
FT /*tag= a
XX
XX WO200259381-A2.
XX
XX PD 01-AUG-2002.
XX
XX PF 07-JAN-2002; 2002WO-US000473.
XX
XX PR 06-JAN-2001; 2001US-0260080P.
XX
XX (GEO) GEN HOSPITAL CORP.
PA
PI Slangenhaupt S, Gusella JF;
XX

PI Slangenaupt S, Gusella JF;
XX WPI; 2002-674806/72.
XX
PT New IKKAP genes with mutations, useful for identifying a subject with
PT familial dysautonomia (FD), or for rapid carrier screening in the
PT Ashkenazi Jewish population, e.g. screening presymptomatic homozygotes or
PT prenatal diagnosis.
XX Claim 1; Page; 109pp; English.
XX
CC The present invention relates to methods and compositions useful for
CC detecting mutations which cause Familial Dysautonomia (FD, Riley-Day
CC syndrome, Hereditary Sensory and Autonomic Neuropathy Type III) [OMIM
CC 223900]. It was found that mutations in the IKKAP gene (see AB080565)
CC are associated with FD. The mutation associated with the major haplotype
CC of FD, FDI mutation, is a base pair (bp) mutation, where the thymine
CC nucleotide located at bp 6 of intron 20 in the IKKAP gene is replaced
CC with a cytosine. This results in skipping of exon 20 in the mRNA from FD
CC patients, although they continue to express varying levels of wild-type
CC message in a tissue-specific manner. The mutation associated with the
CC minor haplotype, FD2 mutation, is a bp mutation, where the guanine
CC nucleotide at bp 2397 (bp 73 of exon 19) is replaced with a cytosine.
CC This bp mutation causes an arginine to proline missense mutation
CC in the IKKAP protein, which is predicted to disrupt a potential
CC phosphorylation site. The IKKAP nucleic acid sequences are useful for
CC identifying a subject with FD and for rapid carrier screening. The IKKAP
CC gene contains 37 exons and maps to chromosome 9q31. Note: the present
CC sequence was not shown in the specification, but was derived from the
CC human wild-type IKKAP sequence given in Fig 6
XX
SQ Sequence 66479 BP; 18271 A; 12400 C; 14128 G; 21680 T; 0 U; 0 Other;
Query Match 72.8%; Score 18.2; DB 6; Length 66479;
Best Local Similarity 87.0%; Pred. No. 5.9e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 AAAAAACATGACCAACTGGGCT 23
Db 60341 AATAAACATGACCAATGGCCT 60319
RESULT 16
AB080565/c
ID ABQ80565 standard; DNA; 66479 BP.
XX
AC ABQ80565;
XX
DT 08-NOV-2002 (first entry)
XX
DE Human IKKAP wild-type gene.
XX
KW Human; IKKAP; Familial Dysautonomia; FD; Riley-Day syndrome;
KW Hereditary Sensory and Autonomic Neuropathy Type III; carrier screening;
KW gene; chromosome 9q31; ds.
XX
OS Homo sapiens.
XX
PN WO200259381-A2.
XX
PD 01-AUG-2002.
XX
PF 07-JAN-2002; 2002WO-US000473.
XX
PR 06-JAN-2001; 2001US-0260080P.
XX
PA (GEO) GEN HOSPITAL CORP.
XX
PI Slangenaupt S, Gusella JF;
XX WPI; 2002-674806/72.
XX
PT New IKKAP genes with mutations, useful for identifying a subject with

PT familial dysautonomia (FD), or for rapid carrier screening in the
PT Ashkenazi Jewish population, e.g. screening presymptomatic homozygotes or
PT prenatal diagnosis.
XX Claim 1; Fig 6; 109pp; English.
XX
CC The present invention relates to methods and compositions useful for
CC detecting mutations which cause Familial Dysautonomia (FD, Riley-Day
CC syndrome, Hereditary Sensory and Autonomic Neuropathy Type III) [OMIM
CC 223900]. It was found that mutations in the IKKAP gene (the present
CC sequence) are associated with FD. The mutation associated with the major
CC haplotype of FD, FDI mutation, is a base pair (bp) mutation, where the
CC thymine nucleotide located at bp 6 of intron 20 in the IKKAP gene is
CC replaced with a cytosine. This results in skipping of exon 20 in the mRNA
CC from FD patients, although they continue to express varying levels of
CC wild-type message in a tissue-specific manner. The mutation associated
CC with the minor haplotype, FD2 mutation, is a bp mutation, where the
CC guanine nucleotide at bp 2397 (bp 73 of exon 19) is replaced with a
CC cytosine. This bp mutation causes an arginine to proline missense
CC mutation (R696P) in the IKKAP protein, which is predicted to disrupt a
CC potential phosphorylation site. The IKKAP nucleic acid sequences are
CC useful for identifying a subject with FD and for rapid carrier screening.
CC The IKKAP gene contains 37 exons and maps to chromosome 9q31
XX
SQ Sequence 66479 BP; 18271 A; 12398 C; 14129 G; 21681 T; 0 U; 0 Other;
Query Match 72.8%; Score 18.2; DB 6; Length 66479;
Best Local Similarity 87.0%; Pred. No. 5.9e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 AAAAAACATGACCAACTGGGCT 23
Db 60341 AATAAACATGACCAATGGCCT 60319
RESULT 17
ABD32804_4
Continuation (5 of 6) of ABD32804 from base 400001 (Mouse cancer-associated genomic DNA
WP Sequence split into 6 fragments LOCUS ABD32804 Accession Abd32804
WP Fragment Name Begin End
WP ABD32804_0 1 110000
WP ABD32804_1 100001 210000
WP ABD32804_2 200001 310000
WP ABD32804_3 300001 410000
WP ABD32804_4 400001 510000
WP ABD32804_5 500001 608916
Query Match 72.8%; Score 18.2; DB 13; Length 110000;
Best Local Similarity 87.0%; Pred. No. 6.3e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 2 AAAAAACATGACCAACTGGGCTG 24
Db 32617 AACGCAATGACCAACTGGCTG 32639
RESULT 18
ABZ16013/c
ID ABZ16013 standard; DNA; 1522 BP.
XX
AC ABZ16013;
XX
DT 21-JAN-2003 (first entry)
XX
DE Arabidopsis thaliana stress regulated gene SEQ ID NO 3818.
XX Arabidopsis thaliana; plant; gene; stress; transgenic; ds.
XX Arabidopsis thaliana.
OS Arabidopsis thaliana.
XX
PN WO200216655-A2.
XX
PD 28-FEB-2002.

XX 24-AUG-2001; 2001WO-US026685.
 XX 24-AUG-2000; 2000US-0227866P.
 PR 26-JAN-2001; 2001US-0264647P.
 PR 22-JUN-2001; 2001US-0300111P.
 XX
 PA (SCRI) SCRIPPS RES INST.
 PA (SYGN) SYNGENTA PARTICIPATIONS AG.
 XX Harper JF, Kreps J, Wang X, Zhu T;
 XX WPI; 2002-304127/34.
 DR
 XX Identifying a stress condition to which a plant cell has been exposed and
 PT producing plants with increased tolerance to these abiotic stresses.
 XX
 XX Claim 144; SEQ ID NO 3818; 577pp + Sequence Listing; English.
 XX
 XX The invention relates to identifying a stress condition to which a plant
 CC cell has been exposed, comprising: (a) contacting nucleic acid
 CC representative of expressed polynucleotides in the plant cell with an
 CC array or probes representative of the plant cell genome; and (b)
 CC detecting a profile of expressed polynucleotides in the plant cell
 CC characteristic of a stress response. The method is useful in the
 CC production of transgenic plants, cells and seeds and in producing plants
 CC with increased tolerance to abiotic stress. The present sequence is that
 CC of an Arabidopsis thaliana stress regulated gene (AB212196-AB217574) used
 CC in methods of the invention. Note: The sequence data for this patent is
 CC not represented in the printed specification but is based on sequence
 CC information supplied to Derwent by the European Patent Office
 XX
 XX Sequence 1522 BP; 392 A; 246 C; 293 G; 591 T; 0 U; 0 Other;
 SQ
 Query Match 72.0%; Score 18; DB 6; Length 1522;
 Best Local Similarity 100.0%; Pred. No. 4e+02;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AAAAAACAATGACCAACT 18
 Db 1386 AAAAAACAATGACCAACT 1369
 RESULT 19
 ADA69122/c
 ID ADA69122 standard; DNA; 1522 BP.
 XX
 AC ADA69122;
 XX
 DT 20-NOV-2003 (first entry)
 DE Arabidopsis thaliana gene, SEQ ID 2445.
 XX
 XX Plant; bacterial infection; fungal infection; viral infection; gene; ds.
 XX
 XX Arabidopsis thaliana.
 OS
 XX Arabidopsis thaliana.
 XX WO2003000898-A1.
 XX
 XX 03-JAN-2003.
 PD
 XX 22-JUN-2001; 2001WO-IB001105.
 PF
 XX 22-JUN-2001; 2001WO-IB001105.
 PR
 XX (SYGN) SYNGENTA PARTICIPATIONS AG.
 PA
 XX Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
 PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;
 XX WPI; 2003-175290/17.
 DR
 XX Identifying at least one gene involved in plant resistance or response to

PT pathogenic infection for conferring resistance or tolerance to a plant to
 PT bacterial, fungal or viral infection by determining or detecting plant
 PT gene expression.
 XX Claim 27; SEQ ID NO 2445; 899pp; English.
 XX
 XX The present invention relates to a method (M1) for identifying genes
 CC involved in plant resistance or response to pathogenic infection. M1
 CC comprises identifying a gene whose expression is significantly altered in
 CC the incompatible interaction of plant gene expression relative to
 CC expression of the gene in an uninfected plant, in a mutant plant that
 CC does not express a gene associated with response to pathogenic infection,
 CC or in a corresponding incompatible or compatible interaction. (M1) is
 CC useful for conferring resistance to resistance or tolerance to a plant to
 CC bacterial, fungal or viral infection. The present sequence was used to
 CC illustrate the invention.
 XX
 XX Sequence 1522 BP; 392 A; 246 C; 293 G; 591 T; 0 U; 0 Other;
 SQ
 Query Match 72.0%; Score 18; DB 8; Length 1522;
 Best Local Similarity 100.0%; Pred. No. 4e+02;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AAAAAACAATGACCAACT 18
 Db 1386 AAAAAACAATGACCAACT 1369
 RESULT 20
 ABZ36201/c
 ID ABZ36201 standard; cDNA; 3824 BP.
 XX
 AC ABZ36201;
 XX
 DT 10-FEB-2003 (first entry)
 DE Human secretory polynucleotide SPTM SEQ ID NO 365.
 XX
 XX Human; SPTM; autoimmune disorder; inflammatory disorder; AIDS; anaemia;
 KW asthma; Crohn's disease; neurological disorder; epilepsy; cancer;
 KW Huntington's disease; Alzheimer's disease; Creutzfeldt-Jakob disease;
 KW multiple sclerosis; Parkinson's disease; cell proliferative disorder;
 KW anti-inflammatory; immunosuppressive; neuroprotective; neurotropic;
 KW neuroleptic; anticonvulsant; cytostatic; antiparkinsonian; anxiolytic;
 KW antipsoriatic; antianaemic; anti-HIV; human immunodeficiency virus;
 KW secretory polynucleotide; secretory protein; gene; ss.
 XX
 OS Homo sapiens.
 XX
 XX WO200283876-A2.
 XX
 XX 24-OCT-2002.
 PD
 XX 27-MAR-2002; 2002WO-US009921.
 PF
 XX 29-MAR-2001; 2001US-0280067P.
 PR 29-MAR-2001; 2001US-0280068P.
 PR 16-MAY-2001; 2001US-0291280P.
 PR 17-MAY-2001; 2001US-0291829P.
 PR 17-MAY-2001; 2001US-0291849P.
 PR 19-JUN-2001; 2001US-0299428P.
 PR 20-JUN-2001; 2001US-0299776P.
 PR 20-JUN-2001; 2001US-0300001P.
 XX
 XX (INCY-) INCYTE GENOMICS INC.
 PA
 XX Daffo A, Jones AL, Tran AB, Dahl CR, Gietzen D, Chinn J;
 XX Dufour GE, Hillman JL, Yu JY, Tuason O, Yap PE, Amshey SR;
 PI Daugherty SC, Dam TC, Liu TF, Nguyen DA, Kleefeld Y, Gerstein EH;
 PI Peralta CH, David MH, Lewis SA, Chen AJ, Panzer SR, Harris B;
 PI Flores V, Marwaha R, Lo A, Lan RY, Urashka ME;
 XX WPI; 2003-075543/07.
 DR

DR P-PSDB; ABP75759.
XX New human secretory proteins and polynucleotides, useful for diagnosing,
PT treating or preventing autoimmune/inflammatory disorders (e.g. AIDS),
PT neurological disorders (e.g. Alzheimer's), or cell proliferations or
PT cancers.
XX Claim 1; SEQ ID NO 365; 459pp + Sequence Listing; English.
XX
XX The invention relates to a secretory polynucleotide (designated sptm)
CC comprising any of 567 polynucleotide sequences (AB235837-AB236403), a
CC naturally occurring polynucleotide sequence at least 90 % identical to
CC the polynucleotide sequence, a polynucleotide complementary to them or an
CC RNA equivalent of them. The polypeptide or polynucleotide are useful for
CC treating, preventing or diagnosing a disease or condition associated with
CC the expression of functional SPTM. These are particularly useful for
CC diagnosing, treating or preventing autoimmune/inflammatory disorders
CC (e.g. acquired immunodeficiency syndrome, anaemia, asthma or Crohn's
CC disease), neurological disorders (e.g. epilepsy, Huntington's disease,
CC dementia, stroke, Alzheimer's disease, Creutzfeldt-Jakob disease,
CC multiple sclerosis, cerebral palsy, Parkinson's disease, anxiety,
CC schizophrenia or amnesia), or cell proliferative disorders (e.g.
CC psoriasis, polycythemia vera, or cancers including adenocarcinoma,
CC leukaemia, lymphoma, melanoma, myeloma, sarcoma or cancers of the brain,
CC breast, cervix or prostate). Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in electronic
CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 3824 BP; 1005 A; 937 C; 934 G; 948 T; 0 U; 0 Other;

Query Match 71.2%; Score 17.8; DB 8; Length 3824;
Best Local Similarity 90.5%; Pred. No. 5.7e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 AACAAATGACCACTGGGCTGT 25
Db 2977 AACAAAGACCACTGGGCTGT 2957
||||| ||||| ||||| ||||| |||||

RESULT 21
ADE28224/c
ID ADE28224 standard; DNA; 5821 BP.
XX ADE28224;
AC
XX
XX 29-JAN-2004 (first entry)
DT
XX Human MDDT DNA - SEQ ID 74.
DE
XX human; MDDT; molecule for disease detection and treatment; cytostatic;
KW antiarteriosclerotic; anticonvulsant; nootropic; neuroprotective;
KW cerebroprotective; anti-HIV; antiallergic; antiinflammatory;
KW thymometric; cell proliferative; cancer; atherosclerosis; neurological;
KW epilepsy; Huntington's disease; stroke; immune; inflammatory; AIDS;
KW allergy; developmental disorder; hypothyroidism; Cushing's syndrome;
KW infection; ds; gene.
XX
XX Homo sapiens.
OS
XX WO2003046152-A2.
PN
XX
XX 05-JUN-2003.
PD
XX
XX 25-NOV-2002; 2002WO-US038446.
PF
XX
XX 28-NOV-2001; 2001US-0334182P.
PR
XX 18-DEC-2001; 2001US-0342052P.
PR
XX 18-JAN-2002; 2002US-0350410P.
PR
XX 01-FEB-2002; 2002US-0353284P.
PR
XX 08-MAR-2002; 2002US-0363649P.
PR
XX (INCY-) INCYTE GENOMICS INC.
PA

PI Tang YT, Chawla NK, Lu DAM, Khan FA, Gandhi AR, Swarnakar A;
PI Azimzai Y, Marquis JP, Sprague WW, Emerling BM, Yue H, Borowsky ML;
PI Becha SD, Ison CH, Elliott VS, Hafalia AJA, Ring HZ, Warren BA;
PI Gietzen KU, Tran UK, Lee SY, Lee EA, Richardson TW, Kable AE;
PI Burford N, Lehr-Mason PM, Gorvad AE, Lee S, Blake JJ, Honchell CD;
PI Thangavelu K, Ramkumar J, Chien D, Jin P, Chang H, Baughn MR;
PI Nguyen DB, Khare R, Bhatia U, Burrill JD, Ho A, Zheng W;
XX WPI; 2003-513643/48.
DR P-PSDB; ADE28155.
XX
XX New human molecules for disease detection and treatment (MDDT), useful
PT for diagnosing, treating and preventing diseases or conditions associated
PT with the aberrant MDDT expression e.g. cancer, AIDS, atherosclerosis,
PT epilepsy, or infections.
XX
XX Claim 5; SEQ ID NO 74; 314pp; English.
XX
XX The invention relates to a novel isolated human MDDT (molecule for
CC disease detection and treatment) polypeptide. The polypeptide of the
CC invention demonstrates cytostatic, antiarteriosclerotic, anticonvulsant,
CC nootropic, neuroprotective, cerebroprotective, anti-HIV, antiallergic,
CC antiinflammatory and thymometric activities and may be useful for
CC diagnosing, treating and preventing a variety of diseases including cell
CC proliferative diseases such as cancer and atherosclerosis, neurological
CC diseases, in particular epilepsy, Huntington's disease and stroke, immune
CC or inflammatory diseases including AIDS and allergies and developmental
CC disorders including hypothyroidism and Cushing's syndrome, as well as
CC infections. The current sequence is that of the human MDDT DNA of the
CC invention.
XX
XX Sequence 5821 BP; 1439 A; 1569 C; 1553 G; 1260 T; 0 U; 0 Other;
SQ
Query Match 71.2%; Score 17.8; DB 10; Length 5821;
Best Local Similarity 90.5%; Pred. No. 6.1e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 AACATGACCACTGGGCTGT 25
Db 5096 AACAAAGACCACTGGGCTGT 5076
||||| ||||| ||||| ||||| |||||

RESULT 22
ADY19422/c
ID ADY19422 standard; DNA; 6225 BP.
XX ADY19422;
AC
XX
XX 05-MAY-2005 (first entry)
DT
XX
XX DNA encoding a PRO polypeptide, SEQ ID NO 5228.
DE
XX Antinflammatory; Immune disorder; Dermatological; Immunosuppressive;
KW Antirheumatic; Antiarthritic; Osteopathic; Hemostatic; Antianemic;
KW Antithyroid; Antidiabetic; Nephrotropic; CNS-Gen.; Hepatotropic;
KW Virucide; Gastrointestinal-Gen.; Antipsoriatic; Antiasthmatic;
KW Antiallergic; ds; gene; diagnosis.
XX
XX Homo sapiens.
OS
XX WO2005016962-A2.
PN
XX
XX 24-FEB-2005.
PD
XX
XX 11-AUG-2004; 2004WO-US026249.
PF
XX
XX 11-AUG-2003; 2003US-0493546P.
PR
XX (GETH) GENENTECH INC.
PA
XX Abbas A, Clark H, Ouyang W, Williams MP, Wood WI, Wu TD;
PI WPI; 2005-182330/19.
DR

XX New nucleic acid encoding PRO polypeptide, useful for diagnosing and
PT treating an immune related disorder, e.g. systemic lupus erythematosus,
PT rheumatoid arthritis, osteoarthritis, thyroiditis, or diabetes mellitus.
XX
XX Claim 1; SEQ ID NO 5228; 159pp; English.
XX
XX The invention relates to an isolated nucleic acid encoding a PRO
CC polypeptide. The polypeptide, agonist or an antagonist, antibody,
CC composition, and method are useful for diagnosing and treating an immune
CC related disorder, e.g. systemic lupus erythematosus, rheumatoid
CC arthritis. The present sequence represents a PRO polypeptide.
XX
XX Sequence 6225 BP; 1629 A; 1534 C; 1562 G; 1500 T; 0 U; 0 Other;
SQ
Query Match 71.2%; Score 17.8; DB 14; Length 6225;
Best Local Similarity 90.5%; Pred. No. 6.2e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 5 AACAAATGACCACTGGGCTGT 25
Db 4359 AACAGGACCACTGGGCTGT 4339
RESULT 23
ABQ61046/c
ID ABQ61046 standard; cDNA; 7028 BP.
XX
XX ABQ61046;
XX
XX 26-FEB-2003 (first entry)
XX
XX Membrane protein MEMAP-32 encoding sequence.
XX
XX Neuroprotective; immunomodulator; cancer; chromosome 11; cytostatic;
KW anti-inflammatory; gene therapy; nutritional supplement; wound; burn;
KW ulcer; Alzheimer's disease; Huntington's disease;
KW amyotrophic lateral sclerosis; autoimmune disorder; inflammation;
KW vulnery; gene; ss.
XX
XX Homo sapiens.
XX
XX WO200231111-A2.
XX
XX 18-APR-2002.
XX
XX 11-OCT-2001; 2001WO-US027760.
XX
XX 12-OCT-2000; 2000US-00687527.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
PI Xue AJ, Yang Y, Wehrman T, Drmanac RT;
XX
XX WPI; 2002-426278/45.
XX
XX N-PSDB; ABP43802.
XX
XX New polypeptides and their encoded proteins, useful as nutritional
PT sources or supplements, or in gene therapy, particularly for treating
PT wounds, Alzheimer's disease, amyotrophic lateral sclerosis, cancer or
PT inflammation.
XX
XX Claim 1; SEQ ID # 259; 357pp + Sequence Listing; English.
XX
XX The invention relates to 446 newly isolated polynucleotide sequences. The
CC activity of polynucleotides of the invention may be described as,
CC inflammatory, neuroprotective, immunomodulator, cytostatic and anti-
CC inflammatory. Compositions comprising nucleic acids of the invention are
CC useful for treating a mammalian subject, or as nutritional sources or
CC supplements. These are useful in gene therapy, particularly for treating
CC wounds, burns or ulcers, Alzheimer's disease, Huntington's disease,
CC amyotrophic lateral sclerosis, autoimmune disorders, cancer or

CC inflammation. The nucleic acids and polypeptides are also useful in
CC diagnostic and research methods. The sequences given in records ABQ60788-
CC ABQ61233 represent polynucleotides of the invention. NOTE: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 7028 BP; 1786 A; 1790 C; 1749 G; 1703 T; 0 U; 0 Other;
SQ
Query Match 71.2%; Score 17.8; DB 6; Length 7028;
Best Local Similarity 90.5%; Pred. No. 6.3e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 5 AACAAATGACCACTGGGCTGT 25
Db 4424 AACAGGACCACTGGGCTGT 4404
RESULT 24
ADQ24033/c
ID ADQ24033 standard; DNA; 7600 BP.
XX
XX ADQ24033;
XX
XX 26-AUG-2004 (first entry)
XX
XX Human soft tissue sarcoma-upregulated DNA - SEQ ID 6853.
XX
XX soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human;
KW ds.
XX
XX Homo sapiens.
XX
XX WO2004048938-A2.
XX
XX 10-JUN-2004.
XX
XX 26-NOV-2003; 2003WO-US038193.
XX
XX 26-NOV-2002; 2002US-0429739P.
XX
XX (PROT-) PROTEIN DESIGN LABS INC.
XX
XX Aziz N, Ginsburg WM, Zlotnik A;
XX
XX WPI; 2004-441208/41.
XX
XX Early detection of soft tissue sarcoma comprises determining expression
PT of a gene in a first soft tissue sample and a normal soft tissue sample
PT and comparing the gene expression, also useful in treating soft tissue
PT sarcoma.
XX
XX Example 2; SEQ ID NO 6853; 210pp; English.
XX
XX The invention relates to a novel method for detecting soft tissue sarcoma
CC which comprises obtaining a first soft tissue sample from an individual
CC and a normal soft tissue sample from the same or different individual,
CC determining the expression of a gene in both samples and comparing the
CC expression of the gene in both soft tissue samples, where a higher level
CC of protein expression in the first soft tissue sample indicates the
CC presence of soft tissue sarcoma. The method of the invention has
CC cytostatic applications and may be useful for detecting soft tissue
CC sarcoma, possibly via gene therapy or vaccine production. The nucleic
CC acid sequences may be useful in diagnostic and screening applications.
CC The current sequence is that of a human soft tissue sarcoma-upregulated
CC DNA of the invention. The current sequence is not shown within the
CC specification per se but was submitted in CD format by the inventor.
XX
XX Sequence 7600 BP; 1939 A; 1919 C; 1880 G; 1829 T; 0 U; 33 Other;
SQ
Query Match 71.2%; Score 17.8; DB 12; Length 7600;
Best Local Similarity 90.5%; Pred. No. 6.4e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
QY      5 AACATGACCACTGGGCTGT 25
DB      4359 AACAGGACCCCACTGGGCTGT 4339

RESULT 25
ACLS5900
ID      ACLS5900 standard; cDNA; 99 BP.
XX      AC
XX      ACLS5900;
XX      24-MAR-2005 (first entry)
XX      Human colon cancer differentially expressed polynucleotide, SEQ ID:3035.
XX      Differential expression; diagnosis; therapy; drug screening; cancer;
KW      neoplasm; colon tumor; breast tumor; pancreas tumor; cytostatic; vaccine;
KW      ss.
XX      Homo sapiens.
XX      WO2005000087-A2.
XX      06-JAN-2005.
XX      13-MAY-2004; 2004WO-US015421.
XX      03-JUN-2003; 2003US-0475872P.
XX      (CHIR ) CHIRON CORP.
XX      Randazzo F, Moler E, Escobedo J, Garcia PD;
XX      WPI; 2005-075421/08.
XX      New isolated polynucleotides, which are differentially expressed in colon
PT      cancer cell, useful for treating cancer, e.g. colon cancer, breast
PT      cancer, or pancreatic cancer.
XX      Claim 1; SEQ ID NO 3035; 97pp; English.
XX      PS
XX      CC The invention relates to 9672 polynucleotides (ACLS3866-ACLS3537) which
XX      are differentially expressed in colon cancer cells. The invention also
XX      relates to vectors and host cells comprising a differentially expressed
XX      polynucleotide of the invention; a method for detecting a cancerous cell
XX      by detection of a gene product of the polynucleotides; a method for
XX      inhibiting a cancerous phenotype of a cell by inhibiting a gene product
XX      of the polynucleotides; a method of treating an individual with cancer by
XX      administration of a modulator of a gene product of the polynucleotides;
XX      and an isolated antibody that specifically binds to a polypeptide encoded
XX      by one of the 9672 polynucleotides. The polynucleotides, polypeptides,
XX      antibodies, and methods are useful for the detection of cancerous cells;
XX      for the diagnosis, prognosis and management of cancer; for the
XX      identification of agents that modulate the phenotype of cancerous cells;
XX      for the identification of therapeutic targets for cancer chemotherapy;
XX      and for the treatment of cancer, especially colon cancer and metastasized
XX      colon cancer, but also breast or pancreatic cancer. The polynucleotides
XX      are also useful as a source of probes or primers for use in diagnostic
XX      methods. The differentially expressed polynucleotides or their encoded
XX      proteins can additionally be used as vaccines to modulate primary immune
XX      responses for the prevention or treatment of cancer. The present sequence
XX      represents a specifically claimed polynucleotide which is differentially
XX      expressed in colon cancer. Note: The sequence data for this patent did
XX      not form part of the printed specification, but was obtained in
XX      electronic format directly from WIPO at
XX      ftp.wipo.int/pub/published_pct_sequences

SQ      Sequence 99 BP; 16 A; 34 C; 17 G; 32 T; 0 U; 0 Other;
Query Match      70.4%; Score 17.6; DB 14; Length 99;
Best Local Similarity 83.3%; Pred. No. 4e+02; Indels 0; Gaps 0;
Matches 20; Conservative 0; Mismatches 4;

RESULT 26
ABX35670/c
ID      ABX35670 standard; cDNA; 397 BP.
XX      AC
XX      ABX35670;
XX      20-FEB-2003 (first entry)
XX      Bovine EST associated with lactation/muscle/fat deposition #835.
XX      Bovine; ss; EST; expressed sequence tag; lactation; LMFD;
KW      muscle deposition; fat deposition; genome mapping; gene identification;
KW      gene analysis; cattle breeding.
XX      Bos Taurus.
XX      OS
XX      US2002137139-A1.
XX      26-SEP-2002.
XX      24-SEP-2001; 2001US-00960352.
XX      12-JAN-1999; 99US-0115707P.
XX      11-JAN-2000; 2000US-00480902.
XX      (BYAT/) BYATT J C.
XX      (MATH/) MATHIALAGAN N.
XX      (TAON/) TAO N.
XX      (WARR/) WARREN W C.
XX      Byatt JC, Mathialagan N, Tao N, Warren WC;
XX      WPI; 2003-110599/10.
XX      New nucleic acid associated with lactation, and muscle and fat
PT      deposition, useful for genome mapping, gene identification and analysis,
PT      cattle breeding, or for genetically improving cattle.
XX      Claim 2; SEQ ID NO 835; 245pp; English.
XX      CC The invention relates to a purified nucleic acid molecule associated with
XX      lactation or muscle and fat deposition (designated LMFD), derived from
XX      cattle, and the LMFD nucleic acid can specifically hybridise to a second
XX      nucleic acid molecule comprising any of 15112 nucleotide sequences,
XX      appearing as ABX34836-ABX49947, or complements of them. Also included are
XX      ; (1) a transformed cell having a nucleic acid comprising an LMFD nucleic
XX      acid linked to a promoter and a 3' non- translated sequence that
XX      functions in the cell to cause termination of transcription and addition
XX      of polyadenylated ribonucleotides to a 3' end of the mRNA molecule; and
XX      (2) determining a level or pattern of a molecule in a bovine cell or
XX      tissue comprising: (a) incubating a marker nucleic acid (comprising any
XX      of the 15112 nucleic acid sequences or its complement or fragment) with a
XX      complementary nucleic acid molecule obtained from the bovine cell or
XX      tissue, where hybridisation between the marker nucleic acid and the
XX      complementary nucleic acid permits the detection of the molecule; and (b)
XX      detecting the level or pattern of the complementary nucleic acid, where
XX      the detection of the complementary nucleic acid is predictive of the
XX      level or pattern of the molecule. The LMFD nucleic acid is used for
XX      determining a level or pattern of a molecule in a bovine cell or tissue.
XX      It is useful for genome mapping, gene identification and analysis, cattle
XX      breeding, preparation of constructs for use in cattle gene expression, or
XX      for genetically improving cattle. The present sequence is one of the
XX      15112 bovine LMFD EST (expressed sequence tag) nucleic acids. Note: The
XX      present sequence was not shown in the specification but was obtained in
XX      electronic format from the USPTO web site:
XX      seqdata.uspto.gov/sequence.html?DocID=20020137139
```


CC	stress condition or for modifying seed oil or protein yield and/or
CC	content. This sequence represents a plant full length insert
CC	polynucleotide that can be used in the recombinant DNA construct of the
CC	invention.
XX	
SQ	Sequence 459 BP; 103 A; 124 C; 158 G; 74 T; 0 U; 0 Other;
	Query Match 70.4%; Score 17.6; DB 13; Length 459;
	Best Local Similarity 83.3%; Pred.No.5.1e+02;
	Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
OY	1 AAAAACAATGACCAACTGGGTG 24
DB	1 ATATAACAAGGCCGACTGGTCTG 24
RESULT 28	
ADX47237	
ID	ADX47237 standard; cDNA; 473 BP.
XX	
AD	ADX47237;
XX	
DT	21-APR-2005 (first entry)
XX	
DE	Plant full length insert polynucleotide seqid 21977.
XX	
KW	plant protectant; plant growth regulant; gene therapy; plant;
KW	recombinant DNA construct; physical array; plant breeding marker;
KW	cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
KW	extreme osmotic condition; pathogen tolerance; pest tolerance;
KW	growth rate; cell cycle pathway; disease resistance;
KW	galactomanan production; lignin production; plant growth regulator;
KW	yield; plant growth; plant development; seed oil; protein yield;
KW	protein content; gene; ss.
OS	Unidentified.
XX	
FN	US2004034888-A1.
XX	
PD	19-FEB-2004.
XX	
PF	28-APR-2003; 2003US-00425114.
XX	
PR	06-MAY-1999; 99US-00304517.
PR	05-NOV-2001; 2001US-00985678.
XX	(LIU//) LIU J.
PA	(ZHOU//) ZHOU Y.
PA	(KOVA//) KOVALIC D K.
PA	(SCRE//) SCREEN S E.
PA	(TAB//) TABASKA J B.
PA	(CAOY//) CAO Y.
XX	
PI	Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
XX	
DR	WPI; 2004-180133/17.
XX	
PT	New recombinant DNA construct, useful for improving plant tolerance to
PT	cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
PT	pests, for conferring increased resistance to plant disease, or for
PT	improving yield.
XX	
PS	Claim 1; SEQ ID NO 21977; 15pp; English.
XX	
CC	The invention describes a recombinant DNA construct comprising a
CC	polynucleotide consisting of a sequence encoding an amino acid sequence
CC	available in electronic form from the US patent office at
CC	ftp.seqsdata.uspo.gov/sequence.html?docID=2004034888. The polynucleotide
CC	of the invention are also useful in physical arrays of molecules and as
CC	plant breeding markers. The recombinant DNA construct is useful for
CC	improving plant tolerance to cold, heat, drought, herbicides, extreme
CC	osmotic conditions, pathogens or pests, for manipulating growth rate in
CC	plant cells by modification of the cell cycle pathway, for conferring

CC increased resistance to plant disease, for producing galactomannan,
CC lignin or plant growth regulators, for increasing the rate of homologous
CC recombination in plants, for improving yield by modification of
CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
CC or by providing improved plant growth and development under at least one
CC stress condition or for modifying seed oil or protein yield and/or
CC content. This sequence represents a plant full length insert
CC polynucleotide that can be used in the recombinant DNA construct of the
CC invention.

XX SQ Sequence 473 BP; 105 A; 131 C; 169 G; 68 T; 0 U; 0 Other;

Query Match 70.4%; Score 17.6; DB 13; Length 473;

Best Local Similarity 83.3%; Pred. No. 5.1e+02;

Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 AAAAAACAATGACCACTGGGCTG 24

Db 444 ATAAACACAGGACCGACTGCTG 467

RESULT 29

AAS27499

ID AAS27499 standard; cDNA; 555 BP.

XX AC AAS27499;

XX DT 07-NOV-2001 (first entry)

XX DE cDNA encoding novel signal transduction pathway protein, Seq ID 534.

XX KW Neuroprotective; cytostatic; dermatological; immunosuppressive; tumour;
KW antiinflammatory; anti-HIV; antibacterial; antiinflammatory; cancer;
KW immune system disorder; rheumatoid arthritis; inflammatory condition;
KW organ transplant rejection; infection; hepatitis C; blood disorder;
KW sickle cell anaemia; hyperproliferative disorder; Gaucher's disease;
KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
KW chromosomal abnormality; Down syndrome; ischaemia; renal disorder;
KW cardiovascular; respiratory; wound healing; endocrine; Addison's disease;
KW reproductive system; gastrointestinal; liver disorder; AIDS; ss;
KW acquired immune deficiency syndrome.

XX OS Homo sapiens.

XX PN WO200154733-A1.

XX PD 02-AUG-2001.

XX PF 17-JAN-2001; 2001WO-US001312.

XX PR 31-JAN-2000; 2000US-0179065P.

PR 04-FEB-2000; 2000US-0180628P.

PR 24-FEB-2000; 2000US-0184664P.

PR 02-MAR-2000; 2000US-0186350P.

PR 16-MAR-2000; 2000US-0189874P.

PR 17-MAR-2000; 2000US-0190076P.

PR 18-APR-2000; 2000US-0198123P.

PR 19-MAY-2000; 2000US-0205515P.

PR 07-JUN-2000; 2000US-0209467P.

PR 28-JUN-2000; 2000US-0214886P.

PR 30-JUN-2000; 2000US-0215135P.

PR 07-JUL-2000; 2000US-0216647P.

PR 07-JUL-2000; 2000US-0216880P.

PR 11-JUL-2000; 2000US-0217487P.

PR 11-JUL-2000; 2000US-0217496P.

PR 14-JUL-2000; 2000US-0218290P.

PR 26-JUL-2000; 2000US-0220963P.

PR 26-JUL-2000; 2000US-0220964P.

PR 14-AUG-2000; 2000US-0224518P.

PR 14-AUG-2000; 2000US-0224519P.

PR 14-AUG-2000; 2000US-0225213P.

PR 14-AUG-2000; 2000US-0225214P.

PR 14-AUG-2000; 2000US-0225266P.

PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234224P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.

XX 05-APR-2001.
 XX 28-SEP-2000; 2000WO-US026524.
 XX 29-SEP-1999; 99US-0157137P.
 XX 03-NOV-1999; 99US-0163280P.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX Ruben SM, Barash SC, Birse CE, Rosen CA;
 XX WPI; 2001-235357/24.
 XX P-PSDB; AAG74146.
 XX Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
 XX useful for preventing, diagnosing and/or treating colorectal cancers.
 XX Claim 1; Page 2666; 9803pp; English.
 XX AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
 XX cancer-associated nucleic acid molecules (N) and proteins (P), where the
 XX proteins are collectively known as colon cancer antigens. The colon
 XX cancer antigens have cytostatic activity and can be used in gene therapy
 XX and vaccine production. N and P may be used in the prevention, diagnosis
 XX and treatment of diseases associated with inappropriate P expression. For
 XX example, N and P may be used to treat disorders associated with decreased
 XX expression by rectifying mutations or deletions in a patient's genome
 XX that affect the activity of P by expressing inactive proteins or to
 XX supplement the patient's own production of P. Additionally, N may be used
 XX to produce the colon cancer-associated Ps, by inserting the nucleic acids
 XX into a host cell and culturing the cell to express the proteins. N and P
 XX can be used in the prevention, diagnosis and treatment of colorectal
 XX carcinomas and cancers. AAH37196 to AAH37204 and AAG77789 represent
 XX sequences used in the exemplification of the present invention. N.B.
 XX Pages 666 to 682 and page 7053 of the sequence listing were missing at
 XX time of publication, meaning no sequences are present for SEQ ID NO:1027
 XX to 1052, 7921 and 7922
 XX Sequence 669 BP; 175 A; 185 C; 142 G; 167 T; 0 U; 0 Other;
 XX
 XX Query Match 70.4%; Score 17.6; DB 4; Length 669;
 XX Best Local Similarity 83.3%; Pred. No. 5.4e+02;
 XX Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 2 AAAAAAATGACCAACTGGGCTGT 25
 DB 140 AAATTCATCAATCACTGCTGT 163
 RESULT 33
 AA215784
 ID AA215784 standard; cDNA; 797 BP.
 XX AA215784;
 XX 12-OCT-1999 (first entry)
 XX Human gene expression product cDNA sequence SEQ ID NO:3253.
 XX Human; gene; gene expression product; diagnosis; therapy; probe;
 XX detection; mapping; tissue typing; profiling; forensic; cancer;
 XX genetic analysis; colorectal cancer; breast cancer; lung cancer; ss.
 XX Homo sapiens.
 XX WO9398972-A2.
 XX 05-AUG-1999.
 XX 28-JAN-1999; 99WO-US001619.
 XX 28-JAN-1998; 98US-0072910P.
 XX

PR 24-FEB-1998; 98US-0075954P.
 PR 31-MAR-1998; 98US-0080114P.
 PR 03-APR-1998; 98US-0080515P.
 PR 03-APR-1998; 98US-0080666P.
 PR 21-OCT-1998; 98US-0105234P.
 PR 28-OCT-1998; 98US-0105877P.
 XX (CHIR) CHIRON CORP.
 XX (HYSE-) HYSEQ INC.
 XX Williams LT, Escobedo J, Innis WA, Garcia PD, Sudduth-Klinger J;
 XX Reinhard C, Giese K, Randazzo F, Kennedy GC, Pot D, Kassam A;
 XX Lamson G, Drmanac R, Crkvenjakov R, Dickson M, Drmanac S, Labat I;
 XX Leshkowitz D, Kita D, Garcia V, Jones WL, Stache-Crain B;
 XX WPI; 1999-494092/41.
 XX Novel human genes and their expression products which are differentially
 XX expressed in different cell types.
 XX Claim 1; Page 1561; 2479pp; English.
 XX The present invention describes a library of human polynucleotides
 XX comprising the sequences given in AA21532 to AA21779. Also described is
 XX a method of detecting differentially expressed genes correlated with the
 XX cancerous state of a mammalian cell, comprising detecting at least one
 XX differentially expressed gene product in a test sample from a cell
 XX suspected of being cancerous, where the gene product is encoded by one of
 XX the 5248 polynucleotide sequences given in AA21532 to AA21779. The
 XX polynucleotides can be used as a source of primers and probes, which can
 XX be used for a variety of purpose, e.g. detection of expression levels,
 XX mapping, tissue typing or profiling, forensics, genetic analysis and
 XX detection of polymorphisms. Polypeptides encoded by the polynucleotides
 XX can be used for raising antibodies for experimental, diagnostic and
 XX therapeutic purposes. The polynucleotides may also be used to construct
 XX arrays for diagnostics (which may be used to determine function of an
 XX encoded protein); and to detect differences in expression levels between
 XX two cells (e.g. to identify abnormal or diseased tissue in a human, to
 XX identify a genetic predisposition or susceptibility to a disease such as
 XX cancer). The polynucleotides of the invention are especially used in the
 XX diagnosis, prognosis and management of colorectal cancer, breast cancer,
 XX and lung cancer. The polynucleotides can also be used to screen for
 XX peptide analogues and antagonists
 XX Sequence 797 BP; 194 A; 217 C; 162 G; 203 T; 0 U; 21 Other;
 XX
 XX Query Match 70.4%; Score 17.6; DB 2; Length 797;
 XX Best Local Similarity 83.3%; Pred. No. 5.6e+02;
 XX Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 2 AAAAAAATGACCAACTGGGCTGT 25
 DB 385 AAATTCATCAATCACTGCTGT 408
 RESULT 34
 ADJ74351/C
 ID ADJ74351 standard; cDNA; 875 BP.
 XX ADJ74351;
 XX 06-MAY-2004 (first entry)
 XX Rat cDNA sequence required for viral infection SeqID 265.
 XX ss; viral infection; tumour suppression; bacterial; parasitic growth;
 XX gene trap; serum survival factor; cytostatic; virucidal; antibacterial;
 XX antiparasitic; rat.
 XX Rattus sp.
 XX WO2004010925-A2.
 XX

```
PD 05-FEB-2004.
XX
XX
PF 02-MAY-2003; 2003WO-US013743.
XX
XX
PR 02-MAY-2002; 2002US-0377136P.
XX
XX
PA (UYVA-) UNIV VANDERBILT.
XX
XX
PI Rubin DH;
XX
XX
DR WPI; 2004-143721/14.
XX
XX
PT Novel gene useful for preparing a composition for reducing or inhibiting
PT viral infection or for suppressing a malignant phenotype in a cell.
XX
XX
PS Claim 1; SEQ ID NO 265; 662pp; English.
XX
XX
CC This invention relates to novel mammalian genes involved in viral
CC infection and tumour suppression. Specifically, it refers to methods for
CC identifying cellular genes that are required for viral, bacterial or
CC parasitic growth, as well as genes used for tumour progression, but that
CC are not essential for cell survival. The present invention describes a
CC 'gene trap' method and screening/ selection process that can isolate
CC those polynucleotides that are associated with a specific process of
CC interest. Furthermore, this method relies on the core discovery that
CC virally infected cells become dependent on a serum survival factor that
CC does not affect non-infected cells. Accordingly, identification of these
CC factors and the appropriate inhibitors thereof, provides a means to treat
CC and/ or prevent viral, bacterial or parasitic growth and infection, as
CC well as tumour growth. The compositions of this invention exhibit
CC cytostatic, virucidal, antibacterial and antiparasitic activities. This
CC polynucleotide is a rat cDNA sequence that is necessary for viral
CC infection, given in an exemplification of the invention.
XX
XX
SQ Sequence 875 BP; 221 A; 210 C; 205 G; 211 T; 0 U; 28 Other;
Query Match 70.4%; Score 17.6; DB 12; Length 875;
Best Local Similarity 83.3%; Pred. No. 5.6e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 AAAAAACATGACCAACTGGGCTG 24
Db 88 AAAAAAATGACACATCTGGGCTG 65
RESULT 35
ADX47961/c
ID ADX47961 standard; cDNA; 1163 BP.
XX
XX
AC ADX47961;
XX
XX
DT 21-APR-2005 (first entry)
XX
XX
DE Plant full length insert polynucleotide seqid 22701.
XX
XX
KW plant protectant; plant growth regulant; gene therapy; plant;
KW recombinant DNA construct; physical array; plant breeding marker;
KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
KW extreme osmotic condition; pathogen tolerance; pest tolerance;
KW growth rate; cell cycle pathway; disease resistance;
KW galactomannan production; lignin production; plant growth regulator;
KW yield; plant growth; plant development; seed oil; protein yield;
KW protein content; gene; ss.
XX
XX
OS Unidentified.
XX
XX
PN US2004034888-A1.
XX
XX
PD 19-FEB-2004.
XX
XX
PF 28-APR-2003; 2003US-00425114.
XX
XX
PR 06-MAY-1999; 99US-00304517.
PR
```

```
PR 05-NOV-2001; 2001US-00985678.
XX
XX
PA (LIUJ/) LIU J.
PA (ZHOU/) ZHOU Y.
PA (KOVA/) KOVALIC D K.
PA (SCRE/) SCREEN S E.
PA (TABA/) TABASKA J E.
PA (CAOY/) CAO Y.
XX
XX
PI Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
XX
XX
DR WPI; 2004-180133/17.
XX
XX
PT New recombinant DNA construct, useful for improving plant tolerance to
PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
PT pests, for conferring increased resistance to plant disease, or for
PT improving yield.
XX
XX
PS Claim 1; SEQ ID NO 22701; 15pp; English.
XX
XX
CC The invention describes a recombinant DNA construct comprising a
CC polynucleotide consisting of a sequence encoding an amino acid sequence
CC available in electronic form from the US patent office at
CC ftp.segdata.uspto.gov/sequence.html?DocID:2004034888. The polynucleotide
CC of the invention are also useful in physical arrays of molecules and as
CC plant breeding markers. The recombinant DNA construct is useful for
CC improving plant tolerance to cold, heat, drought, herbicides, extreme
CC osmotic conditions, pathogens or pests, for manipulating growth rate in
CC plant cells by modification of the cell cycle pathway, for conferring
CC increased resistance to plant disease, for producing galactomannan,
CC lignin or plant growth regulators, for increasing the rate of homologous
CC recombination in plants, for improving yield by modification of
CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
CC or by providing improved plant growth and development under at least one
CC stress condition or for modifying seed oil or protein yield and/or
CC content. This sequence represents a plant full length insert
CC polynucleotide that can be used in the recombinant DNA construct of the
CC invention.
XX
XX
SQ Sequence 1163 BP; 249 A; 318 C; 336 G; 260 T; 0 U; 0 Other;
Query Match 70.4%; Score 17.6; DB 13; Length 1163;
Best Local Similarity 83.3%; Pred. No. 5.9e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 AAAAAACATGACCAACTGGGCTG 24
Db 57 ATAAACAAGGACCGACTGGTCTG 34
RESULT 36
AAD41402
ID AAD41402 standard; cDNA; 1670 BP.
XX
XX
AC AAD41402;
XX
XX
DT 30-OCT-2002 (first entry)
XX
XX
DE Human NZMS-15 cDNA.
XX
XX
KW Human; enzyme; NZMS-15; cell proliferative disorder; hepatitis; cancer;
KW psoriasis; leukaemia; autoimmune disorder; diabetes; arteriosclerosis;
KW acquired immune deficiency syndrome; cardiovascular disorder; anorexia;
KW asthma; hypertension; gastrointestinal disorder; reproductive disorder;
KW gastritis; neurological disorder; epilepsy; pulmonary disorder; AIDS;
KW dementia; embolism; gene therapy; eye disorder; transgenic; gene; ss.
XX
XX
OS Homo sapiens.
XX
XX
Key Location/Qualifiers
FT 1..1512
FT CDS /tag= a
FT /product= "NZMS-15 protein"
```

```
FT sig_peptide 1. .99
FT mat_peptide /*tag= b
FT /*tag= 100. .1509
FT /*tag= c
FT /*tag= c
FT /*product= "Mature NZMS-15 protein"
XX
XX WO200246385-A2.
XX
XX 13-JUN-2002.
XX
XX 04-DEC-2001; 2001WO-US047432.
XX
XX 07-DEC-2000; 2000US-0251824P.
XX
XX 08-DEC-2000; 2000US-0254312P.
XX
XX 14-DEC-2000; 2000US-0255773P.
XX
XX 15-DEC-2000; 2000US-0255940P.
XX
XX 15-DEC-2000; 2000US-0256188P.
XX
XX 21-DEC-2000; 2000US-0257488P.
XX
XX 19-JAN-2001; 2001US-0262839P.
XX
XX 26-JAN-2001; 2001US-0264402P.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Tang YT, Griffin JA, Yue H, Lee EA, Baughn MR, Duggan BM,
PI Walia NK, Lee S, Ramkumar J, Warren BA, Gandhi AR, Lu DM, Lu Y,
PI Yao MG, Ding L, Tribouley CM, Sanjanwala MM, Arvizu C, Hillman JL;
XX
XX WPI; 2002-537565/57.
XX
XX P-PSDB; AAE25391.
XX
XX Novel human enzyme, NZMS useful in diagnosis, prevention or treatment of
PT cell proliferative, autoimmune/inflammatory, cardiovascular,
PT gastrointestinal, neurological, pulmonary, reproductive and eye
PT disorders.
XX
XX Claim 88; Page 170-171; 173pp; English.
XX
XX The invention relates to human enzymes designated NZMS and nucleic acid
CC molecules encoding such proteins. Sequences of the invention are useful
CC for diagnosing, treating or preventing disorders associated with aberrant
CC expression of NZMS. The disorders treated include cell proliferative
CC disorders such as hepatitis, psoriasis, cancer (e.g. leukaemia),
CC autoimmune disorders such as diabetes, acquired immune deficiency
CC syndrome (AIDS), cardiovascular disorders such as arteriosclerosis,
CC hypertension), gastrointestinal disorders (e.g. anorexia, gastritis),
CC neurological disorders (e.g. epilepsy, dementia), pulmonary disorders
CC (e.g. emphysema, asthma), reproductive or eye disorders. Polypeptides of
CC the invention is useful in a number of drug screening techniques and to
CC analyse the proteome of a tissue or cell type. They are also useful as
CC elements on a microarray. Polynucleotides of the invention are useful for
CC creating knockin humanised animals or transgenic animals to model human
CC diseases, in somatic or germline gene therapy, to generate a transcript
CC image of a tissue or cell type, for detecting difference in the
CC chromosomal location due to translocation or inversion among normal,
CC carrier or affected individuals and as hybridisation probes for mapping
CC naturally occurring genomic sequences. The present sequence is human NZMS
CC cDNA
XX
XX Sequence 1670 BP; 460 A; 355 C; 438 G; 417 T; 0 U; 0 Other;
SQ
Query Match 70.4%; Score 17.6; DB 6; Length 1670;
Best Local Similarity 83.3%; Pred. No. 6.2e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Oy 2 AAAAACCAATGATCCAGGGCTGT 25
Db 1032 AAAAACCAATGATCCAGGGCTGT 1055
RESULT 37
AAS27126
ID AAS27126 standard; cDNA; 1921 BP.
XX
```

AC AAS27126;
XX
DT 07-NOV-2001 (first entry)
XX
DE cDNA encoding novel signal transduction pathway protein, Seq ID 161.
XX
KW Neuroprotective; cytostatic; dermatological; immunosuppressive; tumour;
KW antiinflammatory; anti-HIV; antibacterial; antiinflammatory; cancer;
KW immune system disorder; rheumatoid arthritis; inflammatory condition;
KW organ transplant rejection; infection; hepatitis C; blood disorder;
KW sickle cell anaemia; hyperproliferative disorder; Gaucher's disease;
KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
KW chromosomal abnormality; Down syndrome; ischaemia; renal disorder;
KW cardiovascular; respiratory; wound healing; endocrine; Addison's disease;
KW reproductive system; gastrointestinal; liver disorder; AIDS; ss;
XX acquired immune deficiency syndrome.
XX
OS Homo sapiens.
XX
XX WO200154733-A1.
XX
XX 02-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US001312.
XX
XX 31-JAN-2000; 2000US-0179065P.
XX
XX 04-FEB-2000; 2000US-0180628P.
XX
XX 24-FEB-2000; 2000US-0184664P.
XX
XX 02-MAR-2000; 2000US-0186350P.
XX
XX 16-MAR-2000; 2000US-0189874P.
XX
XX 17-MAR-2000; 2000US-0190076P.
XX
XX 18-APR-2000; 2000US-0198123P.
XX
XX 19-MAY-2000; 2000US-0205515P.
XX
XX 07-JUN-2000; 2000US-0209467P.
XX
XX 28-JUN-2000; 2000US-0214886P.
XX
XX 30-JUN-2000; 2000US-0215135P.
XX
XX 07-JUL-2000; 2000US-0216647P.
XX
XX 07-JUL-2000; 2000US-0216880P.
XX
XX 11-JUL-2000; 2000US-0217487P.
XX
XX 14-JUL-2000; 2000US-0217496P.
XX
XX 26-JUL-2000; 2000US-0218290P.
XX
XX 26-JUL-2000; 2000US-0220963P.
XX
XX 14-AUG-2000; 2000US-0220964P.
XX
XX 14-AUG-2000; 2000US-0224518P.
XX
XX 14-AUG-2000; 2000US-0225213P.
XX
XX 14-AUG-2000; 2000US-0225214P.
XX
XX 14-AUG-2000; 2000US-0225266P.
XX
XX 14-AUG-2000; 2000US-0225267P.
XX
XX 14-AUG-2000; 2000US-0225268P.
XX
XX 14-AUG-2000; 2000US-0225270P.
XX
XX 14-AUG-2000; 2000US-0225447P.
XX
XX 14-AUG-2000; 2000US-0225757P.
XX
XX 14-AUG-2000; 2000US-0225758P.
XX
XX 14-AUG-2000; 2000US-0225759P.
XX
XX 18-AUG-2000; 2000US-0226279P.
XX
XX 22-AUG-2000; 2000US-0226681P.
XX
XX 22-AUG-2000; 2000US-0226988P.
XX
XX 22-AUG-2000; 2000US-0227182P.
XX
XX 23-AUG-2000; 2000US-0227009P.
XX
XX 30-AUG-2000; 2000US-0228924P.
XX
XX 01-SEP-2000; 2000US-0229287P.
XX
XX 01-SEP-2000; 2000US-0229343P.
XX
XX 01-SEP-2000; 2000US-0229344P.
XX
XX 01-SEP-2000; 2000US-0229345P.
XX
XX 05-SEP-2000; 2000US-0229509P.
XX
XX 06-SEP-2000; 2000US-0229513P.
XX
XX 06-SEP-2000; 2000US-0230437P.
XX
XX 06-SEP-2000; 2000US-0230438P.
XX
XX 08-SEP-2000; 2000US-0231242P.
XX
XX 08-SEP-2000; 2000US-0231243P.
XX
XX 08-SEP-2000; 2000US-0231244P.
XX
XX 08-SEP-2000; 2000US-0231413P.

PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0233081P.
PR 12-SEP-2000; 2000US-0233968P.
PR 14-SEP-2000; 2000US-0233297P.
PR 14-SEP-2000; 2000US-0233298P.
PR 14-SEP-2000; 2000US-0233399P.
PR 14-SEP-2000; 2000US-023400P.
PR 14-SEP-2000; 2000US-023401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249219P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.

PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX
XX WPI, 2001-465460/50.
XX P-FSDB; AAU17209.
XX
XX Novel polypeptides useful for diagnosing, treating, preventing and/or
XX prognosing disorders related to the proteins, including cancers, immune
XX disorders and neuronal disorders.
XX
XX Claim 1; SEQ ID NO 161; 880pp; English.
XX
XX The invention relates to novel isolated polypeptides (I), and
XX polynucleotides (II). (I), (II) and the antibody to (I) are useful for
XX diagnosing, preventing and treating diseases including immune system
XX disorders (e.g. congenital and acquired immunodeficiencies, autoimmune
XX disorders (e.g. rheumatoid arthritis), inflammatory conditions, organ
XX transplant rejections and graft versus host disease, infectious diseases
XX (e.g. hepatitis C), bleeding disorders, haemoglobin abnormalities and
XX other blood-related disorders (sickle cell anaemia), myeloproliferative
XX disorders, primary haematopoietic disorders, hyperproliferative disorders
XX (e.g. Gaucher's disease and cancer), neurodegenerative disorders (e.g.
XX Alzheimer's disease, Parkinson's disease), chromosomal abnormalities
XX (Down syndrome), ischaemic injury (e.g. stroke), renal disorders (e.g.
XX glomerulonephritis), cardiovascular disorders (e.g. arrhythmia),
XX respiratory disorders, dermatological disorders (e.g. Addison's
XX epithelial cell proliferation, endocrine disorders, in wound healing,
XX disease), reproductive system disorders, gastrointestinal disorder
XX (inflammatory disorders), liver disorders (cirrhosis), as stimulators of
XX B-cell responsiveness to pathogens, activators of T-cells, to induce
XX higher affinity antibodies, and as a means to induce tumour proliferation
XX in pathologies e.g. acquired immune deficiency syndrome (AIDS). AAS26976-
XX AAS27850 represent novel signal transduction pathway protein coding
XX sequences and PCR primers of the invention
XX

Query Match 70.4%; Score 17.6; DB 4; Length 1921;
Best Local Similarity 83.3%; Pred. No. 6.4e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 2 AAAAACAATGACCAACTGGGTGT 25
Db 415 AAAAACAATGATCGCCAGGCTGT 438
RESULT 38
ADB93304
ID ADB93304 standard; cDNA; 1921 BP.
XX
XX ADB93304;
XX
XX 04-DEC-2003 (first entry)
XX
XX Human cDNA encoding a novel protein #151.
XX
XX ss; gene; human; autoimmune disease; Parkinson's disease; silicosis;

KW gastrointestinal disease; atherosclerosis; haemophilia; thrombocytopenia;
KW immunosuppressive agent; adjuvant; enhance immune response;
KW higher affinity antibody induction;
KW increased serum immunoglobulin concentration.
XX

OS Homo sapiens.

XX US2002168711-A1.

XX 14-NOV-2002.

XX 17-JAN-2001; 2001US-0076488.

XX 31-JAN-2000; 2000US-0179065P.

XX 04-FEB-2000; 2000US-0180628P.

XX 28-JUN-2000; 2000US-0214886P.

XX 07-JUL-2000; 2000US-0216647P.

XX 11-JUL-2000; 2000US-0216880P.

XX 11-JUL-2000; 2000US-0217487P.

XX 14-JUL-2000; 2000US-0218290P.

XX 26-JUL-2000; 2000US-0220963P.

XX 14-AUG-2000; 2000US-0224518P.

XX 14-AUG-2000; 2000US-0224518P.

XX 14-AUG-2000; 2000US-0225267P.

XX 14-AUG-2000; 2000US-0225268P.

XX 14-AUG-2000; 2000US-0225270P.

XX 14-AUG-2000; 2000US-0225447P.

XX 14-AUG-2000; 2000US-0225757P.

XX 22-AUG-2000; 2000US-0226868P.

XX 30-AUG-2000; 2000US-0228242P.

XX 01-SEP-2000; 2000US-0229287P.

XX 01-SEP-2000; 2000US-0229343P.

XX 01-SEP-2000; 2000US-0229344P.

XX 05-SEP-2000; 2000US-0229345P.

XX 08-SEP-2000; 2000US-0229509P.

XX 08-SEP-2000; 2000US-0229513P.

XX 21-SEP-2000; 2000US-0234223P.

XX 21-SEP-2000; 2000US-0234274P.

XX 25-SEP-2000; 2000US-0234997P.

XX 27-SEP-2000; 2000US-0235834P.

XX 29-SEP-2000; 2000US-0236327P.

XX 29-SEP-2000; 2000US-0236367P.

XX 29-SEP-2000; 2000US-0236368P.

XX 29-SEP-2000; 2000US-0236369P.

XX 29-SEP-2000; 2000US-0236370P.

XX 02-OCT-2000; 2000US-0236802P.

XX 02-OCT-2000; 2000US-0237037P.

XX 02-OCT-2000; 2000US-0237038P.

XX 02-OCT-2000; 2000US-0237039P.

XX 13-OCT-2000; 2000US-0237040P.

XX 20-OCT-2000; 2000US-0239935P.

XX 20-OCT-2000; 2000US-0240960P.

XX 20-OCT-2000; 2000US-0241785P.

XX 20-OCT-2000; 2000US-0241809P.

XX 01-NOV-2000; 2000US-0244617P.

XX 17-NOV-2000; 2000US-0249299P.

XX 08-DEC-2000; 2000US-0251856P.

XX 08-DEC-2000; 2000US-0251868P.

XX 08-DEC-2000; 2000US-0251869P.

XX (ROSE/) ROSEN C A.

XX (RUBE/) RUBEN S M.

XX (BARA/) BARASH S C.

XX Rosen CA, Ruben SM, Barash SC;

XX WPI; 2003-719985/68.

XX P-PSDB; ADB93917.

PT New isolated polypeptide useful for diagnosing and treating
PT immunosuppressive conditions such as autoimmune disease and Parkinson's
PT disease.
XX

XX Claim 3; SEQ ID NO 161; 345pp; English.

XX The invention relates to an isolated polypeptide. The polypeptide is
CC useful for diagnosing a pathological condition or a susceptibility to a
CC pathological condition in a subject, by determining the presence or
CC amount of expression of the polypeptide in a biological sample and
CC amount of expression of the polypeptide in a biological sample and
CC diagnosing a pathological condition or a susceptibility to a pathological
CC condition based on the presence or amount of expression of the
CC polypeptide. The polypeptide is also useful for identifying a binding
CC partner to the polypeptide, which involves contacting the polypeptide
CC with a binding partner and determining whether the binding partner
CC effects an activity of the polypeptide. The polypeptide or the nucleic
CC acid encoding the polypeptide is useful for preventing, treating, or
CC ameliorating a medical condition, which involves administering the
CC polypeptide or the nucleic acid to a mammalian subject. The nucleic acid
CC is useful for diagnosing a pathological condition or a susceptibility to
CC a pathological condition in a subject, which involves determining the
CC presence or absence of a mutation in the nucleic acid, and diagnosing a
CC pathological condition or susceptibility to a pathological condition
CC based on the presence or absence of the mutation. The polypeptide, the
CC nucleic acid and an antibody to the polypeptide are useful for treating
CC autoimmune disease, Parkinson's disease, silicosis, gastrointestinal
CC disease, atherosclerosis, haemophilia, thrombocytopenia. The polypeptide,
CC the nucleic acid and the antibody are useful as immunosuppressive agents,
CC as adjuvants to enhance immune responses, and as agents to induce higher
CC affinity antibodies and increase serum immunoglobulin concentrations. The
CC present sequence represents cDNA encoding a novel human protein. Note:
CC The sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format direct from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20020168711.

XX SQ Sequence 1921 BP; 598 A; 356 C; 454 G; 513 T; 0 U; 0 Other;

Query Match 70.4%; Score 17.6; DB 10; Length 1921;

Best Local Similarity 83.3%; Pred. No. 6.4e+02;

Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AAAACAATGACCACTGGGCTGT 25

|||||

Db 415 AAAACAATGATCGCCAGGCTGT 438

RESULT 39

AAH79150

ID AAH79150 standard; cDNA; 2229 BP.

XX AC AAH79150;

XX DT 19-NOV-2001 (first entry)

XX DE Human RNA helicase 9 encoding cDNA.

XX Human; RNA helicase 9; cytostatic; virucidal; immunomodulatory;
KW antiinflammatory; haemostatic; malignant neoplasm; HIV; infection;
KW human immunodeficiency virus; immunological disease; cancer;
XX nervous system disease; ss.

OS Homo sapiens.

XX FH Key Location/Qualifiers

XX CDS 800..1042

XX /*tag= a

XX /product= "RNA helicase 9"

XX /note= "claimed in claim 6"

XX WO200166592-A1.

XX PN

XX PD 13-SEP-2001.

XX XX

```
PF 26-FEB-2001; 2001WO-CN000217.
XX
PR 07-MAR-2000; 2000CN-00111929.
XX
XX (BIOW-) BIOWINDOW GENE DEV INC SHANGHAI.
XX
XX MAO Y, XIE Y;
PI
XX WPI; 2001-589932/66.
DR P-PSDB; AAG78128.
XX
XX Human RNA helicase 9 and encoded polynucleotide, for diagnosis and
PT treatment of e.g. nervous system diseases, malignant neoplasm, hemopathy,
PT HIV infection, immunological diseases and various inflammations.
XX
XX Claim 6; Page 29-30; 34pp; Chinese.
XX
XX The invention relates to the human RNA helicase 9 and the encoding
CC polynucleotide with cytostatic, virucidal, immunomodulatory,
CC antiinflammatory and haemostatic activity. The polypeptide and encoded
CC polynucleotide are applicable in diagnosis and treatment of malignant
CC neoplasm, haemopathy, HIV infection, immunological diseases, various
CC inflammations, cancer and nervous system diseases
XX
XX Sequence 2229 BP; 666 A; 459 C; 477 G; 627 T; 0 U; 0 Other;
SQ
Query Match 70.4%; Score 17.6; DB 4; Length 2229;
Best Local Similarity 83.3%; Pred. No. 6.5e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 AAAAAAATGACCAACTGGGCTG 24
DB 1935 AAAAAAAGGTACAACTGGGCTG 1959
RESULT 40
AD035596
ID ADO35596 standard; DNA; 2229 BP.
XX
AC ADO35596;
XX
DT 26-AUG-2004 (first entry)
XX
DE Novel mouse gene sequence #269.
XX
XX mouse; murine; cancer; psoriasis; ulcerative colitis; inflammation;
KW ischaemic heart disease; thrombosis; immune disorder; bacterial disorder;
KW viral disorder; ds; gene.
XX
XX Mus sp.
OS
XX WO2004046310-A2.
XX
XX 03-JUN-2004.
XX
XX 24-OCT-2003; 2003WO-US033948.
XX
XX 15-NOV-2002; 2002US-0426916P.
PR 04-DEC-2002; 2002US-0431158P.
PR 05-DEC-2002; 2002US-0431445P.
PR 05-DEC-2002; 2003US-0431606P.
PR 09-JUN-2003; 2003US-0476621P.
PR 08-JUL-2003; 2003US-0485217P.
PR 08-AUG-2003; 2003US-0485359P.
PR 08-AUG-2003; 2003US-0493332P.
PR 08-AUG-2003; 2003US-0493356P.
XX
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
PA
XX Williams LT, Chu K, Lee E, Hestir K, Hayashizaki Y, Kamiya M;
PI
XX WPI; 2004-431966/40.
XX
XX New mouse nucleic acid molecules and polypeptides, useful for treating
PT cancer, psoriasis, ulcerative colitis, inflammation, ischemic heart
PT disease or thrombosis.
XX
XX Claim 1; SEQ ID NO 269; 263pp; English.
XX
XX The invention comprises 744 novel mouse DNA sequences (genes). The DNA
CC sequences of the invention are useful for treating cancer, psoriasis,
CC ulcerative colitis, inflammation, ischaemic heart disease, thrombosis,
CC immune disorders, bacterial disorders and viral disorders. The present
CC nucleic acid represents a mouse DNA sequence of the invention. NOTE: The
CC present DNA sequence is not shown in the specification, but has been
XX retrieved from the WIPO website.
XX
XX Sequence 2229 BP; 659 A; 418 C; 485 G; 667 T; 0 U; 0 Other;
SQ
Query Match 70.4%; Score 17.6; DB 12; Length 2229;
Best Local Similarity 83.3%; Pred. No. 6.5e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 AAAAAAATGACCAACTGGGCTG 24
DB 341 AAAAAATCAATGCCAACTGGGCAG 364
RESULT 41
ABA20469/c
ID ABA20469 standard; DNA; 2449 BP.
XX
XX ABA20469;
AC
XX 23-JAN-2002 (first entry)
DT
XX Human nervous system related polynucleotide SEQ ID NO 12800.
DE
XX Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;
KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;
KW antiparkinsonian; antisickling; antianaemic; antiarthritic; cancer;
KW antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
KW antiallergic; antidiabetic; antilulcer; anticonvulsant; antifungal;
KW antiparasitic; cardiant; immune disorder; cardiovascular disorder;
KW neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.
XX
XX Homo sapiens.
OS
XX WO200159063-A2.
XX
XX 16-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US001334.
XX
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
```

PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 14-AUG-2000; 2000US-0225799P.
PR 14-AUG-2000; 2000US-0226279P.
PR 14-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226688P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 20-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 25-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0242826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250391P.
PR 01-DEC-2000; 2000US-0251160P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
PI WPI; 2001-541565/60.
XX
XX Nucleic acids encoding 3224 human nervous system antigen polypeptides,
PT useful for preventing, diagnosing and/or treating nervous system cancers
PT and metastases.
XX
XX Disclosure; SEQ ID NO 12800; 1701pp + Sequence Listing; English.
XX
XX The invention relates to novel genes (ABA11004-ABA21534) and proteins
CC (AB414678-AB18001) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in
CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
CC cardiovascular disorders such as myocardial ischaemias; (d) wound healing
CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
CC infectious diseases such as viral, bacterial, fungal and parasitic
CC infections. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX

SQ Sequence 2449 BP; 541 A; 596 C; 600 G; 712 T; 0 U; 0 Other;
Query Match 70.4%; Score 17.6; DB 5; Length 2449;
Best Local Similarity 83.3%; Pred. No. 6.6e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 AAAAAACAATGACCAACTGGGCTG 24
DB 1696 AAGAACCAATGACCAATGACTG 1673
RESULT 42
AAL57079
ID AAL57079 standard; cDNA; 2532 BP.
XX AAL57079;
XX
XX 04-DEC-2003 (first entry)
XX Human sperm generation associated factor 86-90 cDNA sequence.
XX Human sperm generation associated factor 86-90; osteoma; leukaemia; gene;
KW ss.
XX Homo sapiens.
XX
XX Location/Qualifiers
FH Key 44..2416
FT CDS /*tag= a
FT /product= "Human sperm generation associated factor 86-90"
FT
FT
FT
FN CN1381483-A.
XX
XX 27-NOV-2002.
XX
XX 18-APR-2001; 2001CN-00112631.
XX
XX 18-APR-2001; 2001CN-00112631.
XX (BIOW-) BIOWINDOW GENE DEV INC SHANGHAI.
XX Mao Y, Xie Y;
XX WPI; 2003-258236/26.
XX P-PSDB; AAO27292.
XX Polypeptide-human sperm generation associated factor-86.90 and
PT polynucleotide for coding it.
XX
XX Claim 6; Page 24-25; Opp: Chinese.
XX This invention relates to a novel polypeptide, human sperm generation
CC associated factor -86.90. The polypeptide may be used to treat diseases
CC such as osteoma and leukaemia. The present sequence is the cDNA sequence
CC which encodes the human sperm generated associated factor -86.90 of the
CC invention
XX
SQ Sequence 2532 BP; 753 A; 486 C; 634 G; 659 T; 0 U; 0 Other;
Query Match 70.4%; Score 17.6; DB 10; Length 2532;
Best Local Similarity 83.3%; Pred. No. 6.7e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 2 AAAAAACAATGACCAACTGGGCTGT 25
DB 1240 AAAAAACAATGACCAAGGCTGT 1263
RESULT 43
ADE11167
ID ADE11167 standard; cDNA; 2679 BP.
XX

AC ADE11167;
XX 29-JAN-2004 (first entry)
XX Human transporter protein encoding gene SEQ ID NO:1.
XX ss; gene; human; transporter; cytostatic; gene therapy; tumour.
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FH Key 1..2679
FT CDS /*tag= a
FT /product= "transporter protein"
XX
XX WO2003076645-A2.
XX
XX 18-SEP-2003.
XX
XX 05-MAR-2003; 2003WO-US006670.
XX
XX 05-MAR-2002; 2002US-0361343P.
XX (APPL-) APPLERA CORP.
XX Yan X, Yan C, Neelam B;
XX WPI; 2003-748395/70.
XX P-PSDB; ADE11168.
XX New human transporter proteins and nucleic acid molecules, useful for
PT diagnosing, preventing or treating disorders associated with aberrant
PT expression of the transporter proteins, e.g. tumors, or in
PT pharmacogenomic analysis.
XX
XX Claim 4; SEQ ID NO 1; 330pp; English.
XX The invention relates to a novel isolated protein comprising a fully
CC defined sequence of 892 amino acids given in the specification, an
CC allelic variant or an orthologue of the protein and/or a fragment of the
CC protein. A protein of the invention has cytostatic activity, and may have
CC a use in gene therapy. The protein or nucleic acid molecule encoding it
CC is useful as a model for the development of human therapeutic targets,
CC aids in the identification of therapeutic proteins, serve as targets for
CC the development of human therapeutic agents that modulate transporter
CC activity in cells and tissues that express the transporter, or in
CC pharmacogenomic analysis. The protein may be used in substantial and
CC specific assays, in raising antibodies or in eliciting another immune
CC response, as a reagent in assays designed to quantitatively determine the
CC levels of the protein in biological fluids, and as markers for tissues in
CC which the protein is expressed. The nucleic acid molecules may be used as
CC probes, primers, chemical intermediates and in biological assays. These
CC may also be used in diagnosing, preventing or treating disorders
CC associated with human transporter protein, such as tumours. The present
CC sequence encodes the human transporter protein of the invention.
XX
SQ Sequence 2679 BP; 800 A; 517 C; 673 G; 689 T; 0 U; 0 Other;
Query Match 70.4%; Score 17.6; DB 10; Length 2679;
Best Local Similarity 83.3%; Pred. No. 6.7e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 2 AAAAAACAATGACCAACTGGGCTGT 25
DB 1194 AAAAAACAATGACCAAGGCTGT 1217
RESULT 44
ABT16031
ID ABT16031 standard; DNA; 2789 BP.
XX
XX AC ABT16031;
XX

DT 28-MAR-2003 (first entry)
XX NOVX related polynucleotide SEQ ID NO 29.
DE
XX
KW Antidiabetic; anorectic; virucide; antibacterial; fungicide; nootropic;
KW prozooside; neuroprotective; antiparkinsonian; antilipaeamic;
KW NOVX-associated disorder; metabolic disorder; diabetes; anorexia;
KW obesity; infectious disease; cancer-associated cachexia; immune disease;
KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
KW haematopoietic disorder; cancer; dyslipidaemia; metabolic disturbance;
KW neurogenesis; cell differentiation; cell proliferation; haematopoiesis;
KW wound healing; angiogenesis; gene therapy; chromosome mapping;
KW tissue typing; preventive medicine; pharmacogenomic; NOVX; gene; ds.
XX
OS Unidentified.
XX
XX WO200299062-A2.
XX
XX 12-DEC-2002.
XX
XX 04-JUN-2002; 2002WO-US017559.
XX
XX 04-JUN-2001; 2001US-0295607P.
XX 06-JUN-2001; 2001US-0296404P.
XX 06-JUN-2001; 2001US-0296418P.
XX 07-JUN-2001; 2001US-0296575P.
XX 11-JUN-2001; 2001US-0297414P.
XX 12-JUN-2001; 2001US-0297567P.
XX 12-JUN-2001; 2001US-0297573P.
XX 14-JUN-2001; 2001US-0298285P.
XX 15-JUN-2001; 2001US-0298528P.
XX 15-JUN-2001; 2001US-0298556P.
XX 18-JUN-2001; 2001US-0299133P.
XX 19-JUN-2001; 2001US-0299230P.
XX 21-JUN-2001; 2001US-0299949P.
XX 22-JUN-2001; 2001US-0300177P.
XX 28-JUN-2001; 2001US-0301530P.
XX 28-JUN-2001; 2001US-0301550P.
XX 03-JUL-2001; 2001US-0302951P.
XX 12-SEP-2001; 2001US-0318711P.
XX 25-SEP-2001; 2001US-0324687P.
XX 24-OCT-2001; 2001US-0339266P.
XX 16-NOV-2001; 2001US-0337524P.
XX 14-DEC-2001; 2001US-0341143P.
XX 21-FEB-2002; 2002US-0358643P.
XX 21-FEB-2002; 2002US-0359151P.
XX 28-FEB-2002; 2002US-0361195P.
XX 05-MAR-2002; 2002US-0361964P.
XX 10-APR-2002; 2002US-0371346P.
XX 10-APR-2002; 2002US-0371523P.
XX 03-JUN-2002; 2002US-00161493.
XX
XX (CURA-) CURAGEN CORP.
XX
XX Anderson DW, Zerhusen BD, Li L, Zhong M, Casman SJ, Gerlach VL;
XX Shimkets RA, Gorman L, Pena CEA, Kekuda R, Patturajan M, Spytek KA;
XX Leite MW, Rastelli L, Macdougall JR, Taupier RJ, Guo X, Miller CE;
XX Shenoy SG, Hjalt T, Voss EZ, Boldog FL, Malyankar UM, Padigaru M;
XX Ji W, Smithson G, Edinger SR, Millet I, Ellerman K;
XX
XX WPI; 2003-140607/13.
XX P-PSDB; ABJ19332.
XX
XX New isolated NOVX polypeptides and polynucleotides, useful for
XX preventing, diagnosing or treating NOVX-associated disorders, e.g.
XX obesity, cancer, Parkinson's disease, infections, immune disorders, or
XX various dyslipidaemias.
XX
XX Example 12; Page 122; 461pp; English.
XX
XX The invention relates to an isolated polypeptide comprising any of the 36
XX 86-1370 residue amino acid sequences, given in the specification, a
XX mature form of them, or a sequence that is at least 95 % identical to, or

CC having one or more conservative amino acid substitutions in one of the 36
CC amino acid sequences. The polypeptides, nucleic acid molecules and
CC antibodies of the invention are useful in the manufacture of a medicament
CC for treating a syndrome associated with a human disease, preferably a
CC NOVX-associated disorder. The nucleic acid molecules, polypeptides and
CC antibodies are useful for treating, preventing or diagnosing diseases
CC such as metabolic disorders, diabetes, obesity, infectious diseases
CC (viral, bacterial, fungal, helminthic, and protozoal), anorexia, cancer-
CC associated cachexia, neurodegenerative disorders, Alzheimer's disease,
CC Parkinson's disease, immune disorders, haematopoietic disorders, cancer
CC and various dyslipidaemias, or metabolic disturbances associated with
CC obesity, metabolic X syndrome, and wasting disorders. The nucleic acids
CC and polypeptides may also be used as targets for the identification of
CC small molecules that modulate or inhibit e.g. neurogenesis, cell
CC differentiation, cell proliferation, haematopoiesis, wound healing and
CC angiogenesis, in gene therapy, in generation of antibodies that bind
CC immunospecifically to NOVX substances for use in therapeutic or
CC diagnostic methods. The nucleic acids are further used as hybridisation
CC probes, in chromosome mapping, tissue typing, preventive medicine, and
CC pharmacogenomics. This polynucleotide represents a NOVX related DNA
XX sequence of the invention
XX
SQ Sequence 2789 BP; 830 A; 538 C; 695 G; 726 T; 0 U; 0 Other;
Query Match 70.4%; Score 17.6; DB 8; Length 2789;
Best Local Similarity 83.3%; Pred. No. 6.8e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 2 AAAAACATGACCAACTGGCTGT 25
DB 1257 AAAAACATGATCGCCAGGGTGT 1280
RESULT 45
ADO41663
ID ADO41663 standard; cDNA; 2789 BP.
XX
XX ADO41663;
XX
XX 15-JUL-2004 (first entry)
XX
XX Novel human polypeptide NOV12b cDNA.
XX
XX cardiant; antiarteriosclerotic; hypotensive; immunosuppressive;
XX dermatological; anorectic; cytostatic; antidiabetic; haemostatic;
XX anti-HIV; antiasthmatic; antibacterial; virucide; neuroprotective;
XX nootropic; antiparkinsonian; antilipemic; analgesic; antianginal;
XX gene therapy; vaccine; NOVX-associated disorder; cardiomyopathy;
XX atherosclerosis; hypertension; scleroderma; obesity; cancer; diabetes;
XX haemophilia; graft-versus-host disease; AIDS; asthma; Crohn's disease;
XX multiple sclerosis; infection; anorexia; cancer-associated cachexia;
XX neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
XX haematopoietic disorder; dyslipidaemia; wasting disorder;
XX chromosome mapping; tissue typing; preventive medicine; pharmacogenomic;
XX vaccine; human; gene; ss.
XX
XX Homo sapiens.
OS
XX US2004018555-A1.
XX
XX 29-JAN-2004.
XX
XX 03-JUN-2002; 2002US-00161493.
XX
XX 04-JUN-2001; 2001US-0295607P.
XX 06-JUN-2001; 2001US-0296404P.
XX 06-JUN-2001; 2001US-0296418P.
XX 07-JUN-2001; 2001US-0296575P.
XX 11-JUN-2001; 2001US-0297414P.
XX 12-JUN-2001; 2001US-0297567P.
XX 12-JUN-2001; 2001US-0297573P.
XX 14-JUN-2001; 2001US-0298285P.
XX 15-JUN-2001; 2001US-0298528P.

PA (JTW//) JI W.
 PA (SMIT//) SMITHSON G.
 PA (EDIN//) EDINGER S R.
 PA (MILL//) MILLET I.
 PA (ELLE//) ELLERMAN K.
 XX
 PI Anderson DW, Zerhusen BD, Li LE, Zhong M, Casman SJ, Gerlach V;
 PI Shinkets RA, Gorman L, Pena CEA, Kekuda R, Patturajan M, Spytek KA;
 PI Leite MW, Rastelli L, Macdougall JR, Taupier RJ, Guo XS, Miller CE;
 PI Shenoy SG, Hjalte T, Voss EZ, Boldog FL, Malyankar UM, Padigar M;
 PI Ji W, Smithson G, Edinger SR, Millet I, Ellerman K;
 XX
 DR WPI: 2004-122030/12.
 DR P-PSDB; AD041662.
 XX
 XX New NOVX polypeptides and nucleic acids, useful for preventing or
 PT treating NOVX-associated disorders, e.g. cancer, diabetes,
 PT atherosclerosis, asthma or AIDS, and in chromosome mapping, tissue typing
 PT or pharmacogenomics.
 XX
 PS Claim 20; SEQ ID NO 27; 322pp; English.
 XX
 CC The invention describes a new polypeptide comprising: any of the 36 fully
 CC defined amino acid sequences (e.g. 860, 919 or 681 amino acids) given in
 CC the specification; a mature form of (a); a sequence that is at least 95%
 CC identical to (a); or a sequence comprising one or more conservative
 CC substitutions in (a). The polypeptide is useful in the manufacture of a
 CC medicament for treating a syndrome associated with a human disease, the
 CC disease selected from a pathology associated with the polypeptide. These
 CC may also be used in diagnosing, treating or preventing NOVX-associated
 CC disorders such as cardiomyopathy, atherosclerosis, hypertension,
 CC scleroderma, obesity, cancer, diabetes, haemophilia, graft-versus-host
 CC disease, AIDS, asthma, Crohn's disease, multiple sclerosis, infections,
 CC anorexia, cancer-associated cachexia, neurodegenerative disorders (e.g.
 CC Alzheimer's disease or Parkinson's disease), haematopoietic disorders,
 CC dyslipidemias and other wasting disorders associated with chronic
 CC diseases. The nucleic acids are also used as hybridisation probes, in
 CC chromosome mapping, tissue typing, preventive medicine, and
 CC pharmacogenomics. The polypeptides are also useful as vaccines. This
 CC sequence encodes a novel human polypeptide of the invention.
 XX
 SQ Sequence 2812 BP; 824 A; 530 C; 696 G; 762 T; 0 U; 0 Other;
 Query Match 70.4%; Score 17.6; DB 12; Length 2812;
 Best Local Similarity 83.3%; Pred. No. 6.8e+02;
 Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 2 AAAAACAATGACCACTGGGCTGT 25
 Db 1210 AAAAACAATGATCGCCAGGCTGT 1233
 RESULT 48
 ADB62320
 ID ADB62320 standard; cDNA; 3256 BP.
 XX
 AC ADB62320;
 XX
 DT 04-DEC-2003 (first entry)
 XX
 DE Human cDNA encoding clone FCBBF30005360.
 XX
 KW Human; ss; gene; pharmaceutical; diagnostic; gene therapy;
 KW tissue regeneration; cell regeneration; membrane protein;
 KW signal transduction-related protein; transcription-related protein;
 KW osteoporosis; neurological disease; cancer; tumour.
 XX
 OS Homo sapiens.
 XX
 FH Key
 FT CDS
 FT Location/Qualifiers
 FT 29..2710
 FT /*tag= a
 FT /product= "Clone FCBBF30005360 protein"

XX EPI308459-A2.
 XX 07-MAY-2003.
 XX
 XX 28-MAR-2002; 2002EP-00007401.
 XX
 XX 05-NOV-2001; 2001JP-00379298.
 XX 25-JAN-2002; 2002US-00350978.
 XX
 XX (HELI-) HELIX RES INST.
 XX (REAS-) RES ASSOC BIOTECHNOLOGY.
 XX
 XX Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
 XX Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Tamechika I;
 XX Seki N, Yoshikawa T, Otsuka M, Nagahara K, Maeuho Y;
 XX
 XX WPI: 2003-450961/43.
 XX P-PSDB; ADB64290.
 XX
 XX New polynucleotides and polypeptides, useful for developing a diagnostic
 PT marker or medicines for regulation of their expression and activity, or
 PT as targets of gene therapy.
 XX
 PS Claim 1; Page; 222pp; English.
 XX
 CC The invention discloses a polynucleotide comprising a sequence selected
 CC from 1970 fully defined nucleotide sequences which encode novel
 CC polypeptides. Also claimed is a polypeptide encoded by the polynucleotide
 CC or its partial peptide, an antibody binding to the polypeptide or peptide
 CC of the polynucleotide, immunologically assaying the polypeptide or
 CC peptide of the polynucleotide by contacting the polypeptide or peptide
 CC with the antibody of the encoded protein, and observing the binding
 CC between the two, a transformant carrying the polynucleotide in an
 CC expressible manner and an antisense polynucleotide. The oligonucleotide
 CC is useful as a primer for synthesising the polynucleotide, or as a probe
 CC for detecting the polynucleotide. The polynucleotides and encoded
 CC proteins are useful as pharmaceutical agents and many disease-related
 CC genes may be included in them, for developing a diagnostic marker or
 CC medicines for regulation of their expression and activity, or as targets
 CC of gene therapy. The genes are involved in tissue and/or cell
 CC regeneration. Membrane proteins, signal transduction-related proteins,
 CC transcription-related proteins, disease-related proteins and genes
 CC encoding them can be used as indicators for diseases (e.g. osteoporosis,
 CC neurological diseases, cancer, tumours). The cDNA may be used to regulate
 CC the activity or expression of the encoded protein to treat diseases. The
 CC sequence presented is a cDNA of the invention. Note: Some of the sequence
 CC data for this patent is not represented in the printed specification, but
 CC is based on sequence information supplied by the European Patent Office.
 XX
 SQ Sequence 3256 BP; 980 A; 597 C; 777 G; 902 T; 0 U; 0 Other;
 Query Match 70.4%; Score 17.6; DB 10; Length 3256;
 Best Local Similarity 83.3%; Pred. No. 6.9e+02;
 Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 2 AAAAACAATGACCACTGGGCTGT 25
 Db 1225 AAAAACAATGATCGCCAGGCTGT 1248
 RESULT 49
 ACF35897
 ID ACF35897 standard; cDNA; 3318 BP.
 XX
 AC ACF35897;
 XX
 DT 06-NOV-2003 (first entry)
 XX
 DE Human ARP30 polypeptide encoding cDNA.
 XX
 KW ARP; prostate; neoplastic; androgen responsive prostate; ARP15;
 KW cytostatic; gene therapy; human; ARP30; gene; ss.


```
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT CDS 252..2057
XX FT /*tag= a
XX FT /product= "ARP30"
XX PN WO2003060148-A2.
XX XX
XX PD 24-JUL-2003.
XX PF 15-JAN-2003; 2003WO-US001457.
XX PR 15-JAN-2002; 2002US-00053248.
XX XX (SYST-) INST SYSTEMS BIOLOGY.
XX PA Lin B;
XX PI
XX XX WPI; 2003-587287/55.
XX DR P-PSDB; ABR82449.
XX XX Diagnosing or predicting susceptibility to a prostate neoplastic
PT condition by contacting a specimen from the individual with an ARP15
PT binding agent that selectively binds an ARP15 polypeptide.
XX XX Claim 134; Page 210-214; 227pp; English.
XX CC The invention relates to diagnosing or predicting susceptibility to a
CC prostate neoplastic condition. The method involves (a) contacting a
CC specimen from the individual with an androgen responsive prostate
CC specific (ARP)15 binding agent that selectively binds an ARP15
CC polypeptide; (b) determining a test expression level of ARP15 polypeptide
CC in the specimen; and (c) comparing the test expression level to a non-
CC neoplastic control expression level of ARP15 polypeptide, where an
CC altered test expression level as compared to the control expression level
CC indicates the presence of a prostate neoplastic condition in the
CC individual. The method is useful for diagnosing or predicting
CC susceptibility to a prostate neoplastic condition or for treating or
CC reducing severity of a prostate neoplastic condition. The present
CC sequence represents a human ARP30 polypeptide encoding cDNA
XX XX Sequence 3318 BP; 842 A; 987 C; 828 G; 661 T; 0 U; 0 Other;
SQ
Query Match 70.4%; Score 17.6; DB 9; Length 3318;
Best Local Similarity 83.3%; Pred. No. 6.9e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 2 AAAAACCAATGACCAACTGGGCTGT 25
Db 2802 AAATTCAATCACCACCACTGTGCTGT 2825
RESULT 50
ADQ74868
ID ADQ74868 standard; cDNA; 3318 BP.
XX AC
XX AC ADQ74868;
XX DT 07-OCT-2004 (first entry)
XX DE Human androgen responsive prostate specific (ARP) polynucleotide #11.
XX KW Human; androgen responsive prostate specific polynucleotide; ARP; gene;
XX KW ss; prostate neoplastic condition; prostate cancer; cytostatic.
XX OS Homo sapiens.
XX XX US2004137440-A1.
XX PN 15-JUL-2004.
XX PD
```

```
PF 15-JAN-2003; 2003US-00345837.
XX XX
XX PR 15-JAN-2003; 2003US-00345837.
XX PA (LINE/) LIN B.
XX XX
XX PI Lin B;
XX XX
XX DR WPI; 2004-517182/49.
XX DR P-PSDB; ADQ74869.
XX XX New substantially pure androgen responsive specific nucleic acid, useful
PT for diagnosing and treating prostate cancer.
XX XX Claim 148; SEQ ID NO 21; 102pp; English.
XX CC The invention relates to human androgen responsive prostate specific
XX CC (ARP) polynucleotides and the polypeptides they encode. The invention
XX CC also relates to a method of diagnosing or predicting susceptibility to a
XX CC prostate neoplastic condition in an individual and a method for treating
XX CC or reducing the severity of a prostate neoplastic condition in an
XX CC individual. The polynucleotides, polypeptides and methods of the
XX CC invention are useful for diagnosing and treating prostate cancer. This
XX CC sequence represents a human ARP polynucleotide of the invention.
XX XX Sequence 3318 BP; 842 A; 987 C; 828 G; 661 T; 0 U; 0 Other;
SQ
Query Match 70.4%; Score 17.6; DB 12; Length 3318;
Best Local Similarity 83.3%; Pred. No. 6.9e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 2 AAAAACCAATGACCAACTGGGCTGT 25
Db 2802 AAATTCAATCACCACCACTGTGCTGT 2825
Search completed: February 3, 2006, 21:56:33
Job time : 212.111 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 3, 2006, 14:38:25 ; Search time 2951.67 Seconds
(without alignments)
396.277 Million cell updates/sec

Title: US-10-719-900-5
Perfect score: 25
Sequence: 1 aaaaacaatgacaaactgggctgt 25

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues
Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 150 summaries

Database :

- EST:*
- 1: gb_est1:*
 - 2: gb_est2:*
 - 3: gb_est3:*
 - 4: gb_est4:*
 - 5: gb_est5:*
 - 6: gb_est6:*
 - 7: gb_est7:*
 - 8: gb_est8:*
 - 9: gb_est9:*
 - 10: gb_est10:*
 - 11: gb_est11:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	23.4	93.6	327	2	BB204269
2	23.4	93.6	380	1	AI661876
3	23.4	93.6	526	1	AA403507
4	20.8	83.2	340	8	T85622
5	20.8	83.2	468	1	AA826673
6	20.8	83.2	539	6	CA439902
7	20.8	83.2	591	3	BM826284
8	20.8	83.2	727	6	CB241074
9	20.8	83.2	737	6	CD369726
10	20.8	83.2	747	5	BU616130
11	20.8	83.2	789	9	BH722326
12	20.8	83.2	803	2	EG741056
13	20.8	83.2	850	7	CR994650
14	20.8	83.2	1704	4	CR591670
15	20.2	80.8	556	1	AA596588
16	20.2	80.8	562	9	AQ279671
17	20.2	80.8	719	7	CO892131
18	19.8	79.2	373	8	H75316
19	19.8	79.2	451	7	CK650993
20	19.8	79.2	548	7	CK650237
21	19.8	79.2	683	6	CF178286
22	19.4	77.6	422	7	CK011470

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 340)
REFERENCE Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,
AUTHORS Holman, M., Hultman, L., Kucaba, T., Le, M., Lennon, G., Marra, M.,
Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F.,
Travaekis, E., Waterston, R., Williamson, A., Wohldmann, P. and
Wilson, R.
The WashU-Merck EST Project
Unpublished (1995)
JOURNAL Contact: Wilson RK
COMMENT Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Insert Size: 899
High quality sequence stops: 242 Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 899 Std Error: 0.00
Seq primer: M13RP1
High quality sequence stop: 242.
Location/Qualifiers
1..340
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:467848"
/db_xref="taxon:9606"
/clone="IMAGE:112231"
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="DH108 (ampicillin resistant)"
/clone_lib="Soares fetal liver spleen INF1S"
/note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)
with a modified polylinker; Site 1: Pac I; Site 2: Eco RI;
1st strand cDNA was primed with a Pac I - oligo(dT) primer
15' AACTGGAGAAATTAATTAAGATCTTTTCTTTTCTTTT 3',
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Pac I and cloned into the Pac I
and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN
Query Match 83.2%; Score 20.8; DB 8; Length 340;
Best Local Similarity 91.7%; Pred. No. 1.9e+02;
Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 AAAAAACAATGACCAACTGGGCTG 24
|||||
DB 153 AAAAAACAATGTCCAACTGGTCTG 176
|||||

RESULT 5
AA826673/c 468 bp mRNA linear EST 14-APR-1998
LOCUS of34g02.s1 NCI CGAP Kid6 Homo sapiens cDNA clone IMAGE:1423058 3'
DEFINITION similar to contains Alu repetitive element,, mRNA sequence.
AA826673
ACCESSION AA826673.1 GI:2900670
VERSION
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 468)
REFERENCE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
TITLE Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 539)
REFERENCE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
TITLE Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Jose Mercuende
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained
from Dr. M. Bento Soares, bento-soares@uiowa.edu
The following repetitive elements were found in this cDNA
sequence: 1-33, >AT rich#Low complexity (matched complement)
36-164, >ALU (matched complement)
Seq primer: M13 FORWARD
POLYA=yes.
Location/Qualifiers
1..539
/organism="Homo sapiens"
/mol_type="mRNA"

Email: cgapbs-remail.nih.gov
Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: Stratagene, Inc.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 861 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 387.
Location/Qualifiers
1..468
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1423058"
/sex="mixed"
/tissue_type="kidney tumor"
/lab_host="SOLR (kanamycin resistant)"
/clone_lib="NCI CGAP Kid6"
/note="Organ: kidney; Vector: Bluescript SK-; Site 1:
EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer:
Oligo dT. Pooled kidney tumors. 5' adaptor sequence: 5'
CAATTCGGCAGAG 3' 3' adaptor sequence: 5'
CTCAGATTTTTTTTTTTTTTTT 3' Average insert size: 1.0 kb."

ORIGIN
Query Match 83.2%; Score 20.8; DB 1; Length 468;
Best Local Similarity 91.7%; Pred. No. 2e+02;
Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 AAAAAACAATGACCAACTGGGCTG 24
|||||
DB 280 AAAAAACAATGTCCAACTGGTCTG 257
|||||

RESULT 6
CA439902/c 539 bp mRNA linear EST 08-NOV-2002
LOCUS UI-H-D10-aux-k-21-0-UI.s1 NCI CGAP D10 Homo sapiens cDNA clone
DEFINITION UI-H-D10-aux-k-21-0-UI 3', mRNA sequence.
CA439902
ACCESSION CA439902.1 GI:24804322
VERSION
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 539)
REFERENCE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
TITLE Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Jose Mercuende
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained
from Dr. M. Bento Soares, bento-soares@uiowa.edu
The following repetitive elements were found in this cDNA
sequence: 1-33, >AT rich#Low complexity (matched complement)
36-164, >ALU (matched complement)
Seq primer: M13 FORWARD
POLYA=yes.
Location/Qualifiers
1..539
/organism="Homo sapiens"
/mol_type="mRNA"

```

/db_xref="taxon:9606"
/clone="UI-H-D10-aux-k-21-0-UI"
/tissue_type="Lung Focal Fibrosis"
/dev_stage="Adult"
/lab_host="DHI08 (Life Technologies)"
/clone_lib="NCI CGAP D10"
/note="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
NCI CGAP D10 is a cDNA library containing the following
tissue(s): A pool of Lung Focal Fibrosis. The library was
constructed according to Bonaldo, Lennon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA
synthesis was primed with an oligo-dT primer containing a
Not I site. Double stranded cDNA was ligated to an EcoR I
adaptor, digested with Not I, and cloned directionally
into pT7T3-Pac vector. The oligonucleotide used to prime
the synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
(gt)18 tail. The sequence tag for this library is
ATACGCGTC.
TAG_TISSUE=lung with fibrosis
TAG_LIB=UI-H-D10
TAG_SEQ=ATACGCGTC"

ORIGIN
Query Match      83.2%; Score 20.8; DB 6; Length 539;
Best Local Similarity 91.7%; Pred. No. 2e+02;
Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAACAATGACCACTGGGCTG 24
    |||||
Db 337 AAAAAACAATGTCACACTGGTCTG 314
    |||||

RESULT 7
BM826284      591 bp mRNA linear EST 06-MAR-2002
LOCUS K-EST0098334 S22SNU16n1 Homo sapiens cDNA clone S22SNU16n1-106-D09
DEFINITION 5', mRNA sequence.
ACCESSION BM826284
VERSION BM826284.1 GI:19182697
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 591)
AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
TITLE 21C Frontier Korean EST Project 2001
JOURNAL Unpublished (2002)
COMMENT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yonggaung@mail.kribb.re.kr
Plate: 106 row: D column: 09
High quality sequence stop: 591.
Location/Qualifiers
1. .591
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S22SNU16n1-106-D09"
/sex="F"
/tissue_type="Ascites"
/cell_type="Lymphoblast-like"
/cell_line="SNU-16"
/lab_host="DHI08"

FEATURES
source
1. .727
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-CF-FN0-afx-m-14-0-UI"
/tissue_type="Human Lung Epithelial cells"
/lab_host="DHI08 (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-CF-FN0"
/note="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
UI-CF-FN0 is a subtracted cDNA library derived from two
normalized Human lung epithelial cell libraries (SNI and
DUI) The library was subtracted according to according to

```

Ronaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. For additional information, contact:
bento-soares@uiowa.edu
TAG TISSUE=Human Lung Epithelial Cell Lines untreated LPS
6hr to LPS 24h
TAG LIB=UI-CF-PN0
TAG_SEQ=CTGCTCAGGT"

ORIGIN

Query Match 83.2%; Score 20.8; DB 6; Length 727;
Best Local Similarity 91.7%; Pred. No. 2.1e+02;
Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AAAAAACAATGACCAACTGGGCTG 24
||||| ||||| ||||| ||||| |||||
Db 645 AAAAAACAATGTCCTCAACTGGTCTG 622

RESULT 9

CD369726/c
LOCUS
DEFINITION UI-H-Ft1-bju-b-05-0-UI.s1 NCI CGAP Ft1 Homo sapiens cDNA clone
UI-H-Ft1-bju-b-05-0-UI 3', mRNA sequence.

ACCESSION CD369726
VERSION CD369726.1 GI:31153816

KEYWORDS

SOURCE EST.

ORGANISM

Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE

1 (bases 1 to 737)

NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: Dr. Gary W. Hunninghake, U of I
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
<http://genome.uiowa.edu/distribution/cgap.html>

The following repetitive elements were found in this cDNA

sequence: 314-336, >AT rich#Low complexity (matched complement)

341-469, >ALU (matched complement)

Seq primer: M13 FORWARD

POLYA=Yes.

FEATURES

source

Location/Qualifiers

1..737
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-H-Ft1-bju-b-05-0-UI"
/tissue_type="Adult"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI CGAP Ft1"
/note="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site_1: EcoR I; Site_2: Not I;
NCI CGAP Ft1 is a normalized cDNA library constructed from a pool of 81 RNA samples from Alveolar Macrophages challenged with different treatments. The mRNA samples were a mixture of these conditions (times refer to incubations following isolation by bronchoalveolar lavage) (some normal donor macrophages were cultured in some of the conditions, other donor macrophages in different conditions). The mRNA samples were pooled for library construction. Control 0 hours; control 3 hours; control 24 hours; LPS 100 ng/ml, 3 hours; LPS 100 ng/ml, 24 hours; PMA 10 ng/ml, 3 hours; PMA 10 ng/ml, 24 hours; Klebsiella moi 10, 3 hours; Klebsiella moi 10, 24 hours; Staph aureus

moi 10, 3 hours; Staph aureus moi 10, 24 hours; Adenoviral vector (Ad5 CMV eGFP), moi 500, 3 hours; Adenoviral vector (Ad5 CMV eGFP), moi 500, 24 hours; wt adenovirus moi 500, 3 hours; wt adenovirus moi 500, 24 hours; Ad vector + LPS 3 hours; Ad vector + LPS 24 hours; wt adenovirus + LPS 3 hours; wt adenovirus + LPS 24 hours. The library was normalized according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (drr)18 tail. The sequence tag for this library is GGCATGCCG. The tissue was provided by Dr. Gary W. Hunninghake of the University of Iowa.

TAG TISSUE=Human Lung Alveolar Macrophage

TAG LIB=UI-H-Ft1

TAG_SEQ=GGCATGCCG"

ORIGIN

Query Match 83.2%; Score 20.8; DB 6; Length 737;
Best Local Similarity 91.7%; Pred. No. 2.1e+02;
Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AAAAAACAATGACCAACTGGGCTG 24
||||| ||||| ||||| ||||| |||||
Db 643 AAAAAACAATGTCCTCAACTGGTCTG 620

Qy 1 AAAAAACAATGACCAACTGGGCTG 24

||||| ||||| ||||| ||||| |||||

Db 643 AAAAAACAATGTCCTCAACTGGTCTG 620

RESULT 10

BU616130/c

LOCUS

DEFINITION UI-H-DF0-bex-b-09-0-UI.s1 NCI CGAP DF0 Homo sapiens cDNA clone

UI-H-DF0-bex-b-09-0-UI 3', mRNA sequence.

ACCESSION BU616130

VERSION BU616130.1 GI:23282338

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

1 (bases 1 to 747)

NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: Dr. Jose Mercuende

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Clone distribution information can be obtained from Dr. M. Bento Soares, bento-soares@uiowa.edu

The following repetitive elements were found in this cDNA

sequence: 1-25, >POLY A#Simple repeat (matched complement) 321-347,

>AT rich#Low complexity 350-478, >ALU (matched complement)

Seq primer: M13 FORWARD

POLYA=Yes.

FEATURES

source

Location/Qualifiers

1..747
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-H-DF0-bex-b-09-0-UI"
/tissue_type="Subchondral Bone"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI CGAP DF0"
/note="Organ: Bone; Vector: pT7T3-Pac (Pharmacia) with a

modified polylinker; Site 1: EcoR I; Site 2: Not I;
 NCI CGAP.DF0 is a cDNA library containing the following
 tissue(s): Subchondral Bone. The library was constructed
 according to Bonaldo, Lennon and Soares, Genome Research,
 6:791-806, 1996. First strand cDNA synthesis was primed
 with an oligo-dT primer containing a Not I site. Double
 stranded cDNA was ligated to an EcoR I adaptor, digested
 with Not I, and cloned directionally into pTVT3-Pac
 vector. The oligonucleotide used to prime the synthesis of
 first-strand cDNA contains a library tag sequence that is
 located between the Not I site and the (dT)18 tail. The
 sequence tag for this library is GTTAAGCGTC.
 TAG TISSUE=subchondral bone
 TAG_LiB=UI-H-DF0
 TAG_SEQ=GTTAAGCGTC"

ORIGIN

Query Match 83.2%; Score 20.8; DB 5; Length 747;
 Best Local Similarity 91.7%; Pred. No. 2.1e+02;
 Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAACAATGACCAACTGGGCTG 24

Db 651 AAAAAACAATGTCCAACTGGTCTG 628

RESULT 11

BH722326/c
 LOCUS BH722326 789 bp DNA linear GSS 20-FEB-2002
 DEFINITION BOMGS07TR BO_2_3_KB Brassica oleracea genomic clone BOMGS07,
 genomic survey sequence.

ACCESSION BH722326

VERSION BH722326.1 GI:18825399

KEYWORDS GSS.

SOURCE Brassica oleracea

ORGANISM Brassica oleracea

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Eukaryota; Magnoliophyta; eudicotyledons; core eudicotyledons;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 rosids; eurosids II; Brassicales; Brassicaceae; Brassica.

AUTHORS 1 (bases 1 to 789)

AYELE, M., HAAS, B. J., KUMAR, N., WU, H., XIAO, Y., VAN AKEN, S.,
 ULLERBACK, T. R., WORTMAN, J. R., WHITE, O. R. and TOWN, C. D.

Whble genome shotgun sequencing of Brassica oleracea and its
 application to gene discovery and annotation in Arabidopsis

JOURNAL Genome Res. 15 (4), 487-495 (2005)

PUBMED 15805490

COMMENT Other GSSs: BOMGS07TF

Contact: Chris Town

TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA.

Tel: 301-838-3523

Fax: 301-838-0208

Email: cdtown@tigr.org

DNA is from a doubled haploid provided by Tom Osborn.

Seq primer: TR

Class: sheared ends.

Location/Qualifiers

1..789

/organism="Brassica oleracea"

/mol_type="genomic DNA"

/strain="TO100DH3"

/db_xref="taxon:3712"

/clone_lib="BOMGS07"

/notes="Vector: pBOS1; Site 1: BstXI; 2-3 kb sheared

genomic DNA inserted into pBOS1 using BstXI linkers"

ORIGIN

Query Match 83.2%; Score 20.8; DB 9; Length 789;
 Best Local Similarity 91.7%; Pred. No. 2.1e+02;
 Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAACAATGACCAACTGGGCTG 24

Db 392 AAAAAACAATGACCAACTGGGTTG 369

RESULT 12

LOCUS BG741056

DEFINITION

ACCESSION BG741056

VERSION BG741056.1 GI:14051709

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 803)

AUTHORS NIH-MGC http://mgc.nci.nih.gov/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: James Cleaver, M.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA

Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM10637 row: b column: 23

High quality sequence stop: 697.

Location/Qualifiers

1..803

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:4779574"

/lab_host="DH10B (T1 phage-resistant)"

/clone_lib="NCI CGAP_Skn3"

/note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;

Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.

Average insert size 1.5kb. Library constructed by Life

Technologies. Note: this is a NCI CGAP Library."

ORIGIN

Query Match 83.2%; Score 20.8; DB 2; Length 803;

Best Local Similarity 91.7%; Pred. No. 2.1e+02;

Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAACAATGACCAACTGGGCTG 24

Db 406 AAAAAACAATGTCCAACTGGTCTG 429

RESULT 13

LOCUS CR994650

DEFINITION

ACCESSION CR994650

VERSION CR994650.1 GI:68288535

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 850)

AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

Heil, O., Ebert, L., Hennig, S., Henze, S., Radelof, U., Schneider, D.

Human T-Lymphocytes library

Unpublished (2005)

JOURNAL

COMMENT Contact: Inge Arlart
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Email: www.rzpd.de
RZPD; RZPDp9016M1729.
RZPDLIB; (Human T-Lymphocytes) RZPD LIB No.9016
http://www.rzpd.de/cgi-bin/products/set.cgi?libNo=9016 Contact:
Inge Arlart
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Tel: +49 30 32639 100
Fax: +49 30 32639 111
www.rzpd.de
This clone is available from RZPD:
http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneID=RZPDp9016M1729
contact RZPD (product- support@rzpd.de) for further information.
Primer name: q3.4 , Primer sequence: CGGATAACAATTTCACACAG.
FEATURES
Location/Qualifiers
1. 850
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="RZPDp9016M1729"
/tissue_type="T-Lymphocytes"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="RZPD no.9016"
/note="Vector: pQE80LSN_cloned; Site_1: SalI; Site_2:
NotI; vector:
http://www.rzpd.de/info/vectors/pQE80LSN_cloned.pic.shtml
; 1st strand cDNA was prepared from mRNA obtained from
human T-Lymphocytes with a NotI - oligo(dT) primer [5'
GACTAGTCTAGATCGGAGCGCCCTTTTCTTTT 3'].
Double-stranded cDNA was ligated to SalI adaptors,
digested with NotI and cloned into the NotI and SalI sites
of the pQE80LSN_cloned vector"

ORIGIN
Query Match 83.2%; Score 20.8; DB 7; Length 850;
Best Local Similarity 91.7%; Pred. No. 2.1e+02;
Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 AAAAACAATGACCACTGGGCTG 24
|||||
Db 363 AAAAACAATGTCACACTGGTCTG 386
|||||

CRS91670 1704 bp mRNA linear HTC 21-JUL-2004
full-length cDNA clone CS0DF030YN16 of Fetal brain of Homo sapiens
(human).
ACCESSION CRS91670
VERSION CRS91670.1 GI:50472477
KEYWORDS HTC; cNSLT cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE 1 (bases 1 to 1704)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue
2 (bases 1 to 1704)
Genoscope.
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)

COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
Location/Qualifiers
1. 1704
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DF030YN16"
/tissue_type="Fetal brain"
/plasmid="pCMVSPORT_6"

ORIGIN
Query Match 83.2%; Score 20.8; DB 4; Length 1704;
Best Local Similarity 91.7%; Pred. No. 2.3e+02;
Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 AAAAACAATGACCACTGGGCTG 24
|||||
Db 283 AAAAACAATGTCACACTGGTCTG 260
|||||

RESULT 15
AA596588
LOCUS AA596588
DEFINITION vms8f12.r1 Stratagene mouse Tcell 937311 Mus musculus cDNA clone
IMAGE:1002479 5', mRNA sequence.
ACCESSION AA596588
VERSION AA596588.1 GI:2412023
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 556)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wyllie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
TITLE The WashU-HMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:566695
Seq primer: -28ml3 rev1 ET from Amersham
High quality sequence stop: 442.
Location/Qualifiers
1. 556
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:1002479"
/tissue_type="Tcell"
/dev_stage="M30 CD4+ cells"
/lab_host="SOLR (kanamycin resistant)"
/clone_lib="Stratagene mouse Tcell 937311"
/note="Organ: blood; Vector: pBluescript SK-; Site 1:
EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer:
Oligo dT. M30 CD4+ cells. Average insert size: 1.0 kb;
Uni-ZAP XR Vector; ~5' adaptor sequence: 5' GAATTCGCACGAG
3' -3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3"

FEATURES
Location/Qualifiers
1. 556
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:1002479"
/tissue_type="Tcell"
/dev_stage="M30 CD4+ cells"
/lab_host="SOLR (kanamycin resistant)"
/clone_lib="Stratagene mouse Tcell 937311"
/note="Organ: blood; Vector: pBluescript SK-; Site 1:
EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer:
Oligo dT. M30 CD4+ cells. Average insert size: 1.0 kb;
Uni-ZAP XR Vector; ~5' adaptor sequence: 5' GAATTCGCACGAG
3' -3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3"

ORIGIN

```

Query Match      80.8%; Score 20.2; DB 1; Length 556;
Best Local Similarity 88.0%; Pred. No. 3.7e+02;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAACAATGACCAACTGGGCTGT 25
   ||||| ||||| ||||| |||||
Db 373 AAGAAACAAGACCAACTGAGCTGT 397

RESULT 16
AQ279671/c
LOCUS
DEFINITION
CITBI-E1-2513L17-TR CITBI-E1 Homo sapiens genomic clone 2513L17,
genomic survey sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
1 (bases 1 to 562)
Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and
Venter,J.C.
Use of a random human BAC End Sequence Database for Sequence-Ready
Map Building
Unpublished (1998)
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@igir.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
Seq primer: M13 Reverse
Class: BAC ends.

FEATURES
source
Location/Qualifiers
1..562
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="2513L17"
/sex="male"
/cell_type="sperm"
/clone_lib="CITBI-E1"
/notes="Vector: pBelOBAC11; Site_1: EcoRI; Site_2: EcoRI;
Caltech Human BAC Library D"

ORIGIN
:
Query Match      80.8%; Score 20.2; DB 9; Length 562;
Best Local Similarity 88.0%; Pred. No. 3.7e+02;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAACAATGACCAACTGGGCTGT 25
   ||||| ||||| ||||| |||||
Db 435 AAAAAACAATGACTATCTGGGCTGT 411

RESULT 17
CO892131/c
LOCUS
DEFINITION
BovGen_20456 normal cattle brain Bos taurus cDNA clone
RZPDp1056D2337Q 5', mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
Bos taurus (cow)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
1 (bases 1 to 719)
Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and
Venter,J.C.
Use of a random human BAC End Sequence Database for Sequence-Ready
Map Building
Unpublished (1998)
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@igir.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
Seq primer: M13 Reverse
Class: BAC ends.

FEATURES
source
Location/Qualifiers
1..719
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/clone="RZPDp1056D2337Q"
/sex="female"
/tissue_type="brain tissue"
/dev_stage="adult brain"
/clone_lib="normal cattle brain"
/notes="Organ: brain; Vector: pSport1; Site 1: NotI;
Site 2: Sali; Random primed and directionally cloned in
pSport1 vector using NotI
(5'-pGATAGTCTAGATCGGACGGCGGCC (T)15-3' and Sali 5'-
TCGACCCACGCGTCCG-3' adapters (Gibco BRL)"

ORIGIN
Query Match      80.8%; Score 20.2; DB 7; Length 719;
Best Local Similarity 88.0%; Pred. No. 3.8e+02;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAACAATGACCAACTGGGCTGT 25
   ||||| ||||| ||||| |||||
Db 162 AAAAAACAATGAATACTGTGCTGT 138

RESULT 18
H75316
LOCUS
DEFINITION
yu06h02.r1 Soares fetal liver spleen INFLS Homo sapiens cDNA clone
IMAGE:233043 5', mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
1 (bases 1 to 373)
Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B.,
Chissoe,S., Dietrich,N., Dubuque,T., Favello,A., Gish,W.,
Hawkins,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N.,
Mardis,E., Moore,B., Morris,M., Parsons,J., Prange,C., Rifkin,L.,
Rohlfing,T., Schellenberg,K., Soares,M.B., Tan,F., Thierry-Mieg,J.,
Trevaskis,E., Underwood,K., Wohldmann,P., Waterston,R., Wilson,R.

```

```

ORGANISM
Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
Pecora; Bovidae; Bovinae; Bos.
REFERENCE
1 (bases 1 to 719)
Hennig,S., Janitz,M., Herwig,R. and Williams,J.
Generation, annotation, evolutionary analysis and database
integration of 14969 cattle EST clusters
Unpublished (2004)
Contact: Hennig S
laboraty 123, dept.Lehrach
Max-Planck-Institut fuer Molekulare Genetik
Innestr.63-73, D-14195 Berlin, Germany
Tel: +49 30 8413 1612
Fax: +49 30 8413 1380
Email: hennig@molgen.mpg.de
The library was characterised by oligonucleotide fingerprinting
(ONFP) to reduce sequencing redundancy. According to the ONFP
procedure, clones that display the same hybridisation matrix with a
battery of 200 8mer oligonucleotides are grouped into clusters. One
clone per ONFP cluster was selected for sequencing. cDNA clones and
filters are distributed via Deutsches Ressourcenzentrum fuer
Genomforschung GmbH (http://www.rzpd.de).
PCR Primers
FORWARD: 5' CCCCGAGCTTTACACTTTATGTCGCGCTCG 3' (M13RSP) 5'-seq
BACKWARD: 5' GCTATTACGCGAGCTGGCGAAAGGGGATGTG 3' (M13FSP) 3'-seq
Seq primer: 5'-CCGTCGCGAATTCGCGGT-3' (M13RSP).
FEATURES
source
Location/Qualifiers
1..719
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/clone="RZPDp1056D2337Q"
/sex="female"
/tissue_type="brain tissue"
/dev_stage="adult brain"
/clone_lib="normal cattle brain"
/notes="Organ: brain; Vector: pSport1; Site 1: NotI;
Site 2: Sali; Random primed and directionally cloned in
pSport1 vector using NotI
(5'-pGATAGTCTAGATCGGACGGCGGCC (T)15-3' and Sali 5'-
TCGACCCACGCGTCCG-3' adapters (Gibco BRL)"

```

TITLE
JOURNAL
PUBMED
COMMENT

and Maria,M.
Generation and analysis of 280,000 human expressed sequence tags
Genome Res. 6 (9), 807-828 (1996)
889549
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Insert Size: 857
High quality sequence stops: 313
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 857 Std Error: 0.00
Seq primer: M13RP1
High quality sequence stop: 313.

Location/Qualifiers
source
1..373
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:3786220"
/db_xref="taxon:9606"
/clone="IMAGE:233043"
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares fetal liver spleen INFLS"
/note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)
with a modified polylinker; Site 1: Pac I; Site 2: Eco RI;
1st strand cDNA was primed with a Pac I - oligo(dT) primer
15' AACGTGAGAAATTAATTAAGATCTTTTCTTTTCTTTT 3',
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Pac I and cloned into the Pac I
and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M.Fatima Bonaldo."

FEATURES

Query Match 79.2%; Score 19.8; DB 8; Length 373;
Best Local Similarity 91.3%; Pred. No. 5.4e+02;
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AAAACAATGACCAACTGGCTG 24
|||||
Db 146 AAAACAATGCAACTGCTG 168
|||||

RESULT 19
CK650993
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

Manihot esculenta (cassava)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids I; Malpighiales; Euphorbiales; Crotonoideae;
Manihoteae; Manihot.
1 (bases 1 to 451)
Lopez,C., Jorge,V., Piegue,B., Mba,C., Cortes,D., Restrepo,S.,
Soto,M., Laudie,M., Berger,C., Cooke,R., Delsey,M., Tohme,J. and
Verdier,V.
A unigene catalogue of 5700 expressed genes in cassava
Plant Mol. Biol. 56 (4), 541-554 (2004)

ORIGIN

Query Match 79.2%; Score 19.8; DB 8; Length 373;
Best Local Similarity 91.3%; Pred. No. 5.4e+02;
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AAAACAATGACCAACTGGCTG 24
|||||
Db 146 AAAACAATGCAACTGCTG 168
|||||

RESULT 19
CK650993
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

Manihot esculenta (cassava)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids I; Malpighiales; Euphorbiales; Crotonoideae;
Manihoteae; Manihot.
1 (bases 1 to 451)
Lopez,C., Jorge,V., Piegue,B., Mba,C., Cortes,D., Restrepo,S.,
Soto,M., Laudie,M., Berger,C., Cooke,R., Delsey,M., Tohme,J. and
Verdier,V.
A unigene catalogue of 5700 expressed genes in cassava
Plant Mol. Biol. 56 (4), 541-554 (2004)

REFERENCE

AUTHORS
Lopez,C., Jorge,V., Piegue,B., Mba,C., Cortes,D., Restrepo,S.,
Soto,M., Laudie,M., Berger,C., Cooke,R., Delsey,M., Tohme,J. and
Verdier,V.
A unigene catalogue of 5700 expressed genes in cassava
Plant Mol. Biol. 56 (4), 541-554 (2004)
Contact: Verdier V
Laboratoire genome et developpement des plantes, UMR5096
IRD-CNRS-Universite de Perpignan-CIAT

52 AV Paul Alduy Perpignan 66860 France
Tel: (33) 4 68 66 17 74
Fax: (33) 4 68 66 84 99
Email: Valerie.Verdier@ird.fr
Seq primer: T3.
Location/Qualifiers
source
1..451
/organism="Manihot esculenta"
/mol_type="mRNA"
/cultivar="MPer183"
/db_xref="taxon:3983"
/clone_lib="MPer183 cassava lambda zap"
/note="Organ: Root; Vector: pBluescript SK (+/-); Site_1:
EcoRI; Site_2: XhoI; Stratagene lambda Zap Library"

FEATURES

Query Match 79.2%; Score 19.8; DB 7; Length 451;
Best Local Similarity 91.3%; Pred. No. 5.5e+02;
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AAAACAATGACCAACTGGCTG 24
|||||
Db 300 AAAACAATGACCAACTGAGCTG 322
|||||

RESULT 20
CK650237
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

Manihot esculenta (cassava)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids I; Malpighiales; Euphorbiales; Crotonoideae;
Manihoteae; Manihot.
1 (bases 1 to 548)
Lopez,C., Jorge,V., Piegue,B., Mba,C., Cortes,D., Restrepo,S.,
Soto,M., Laudie,M., Berger,C., Cooke,R., Delsey,M., Tohme,J. and
Verdier,V.
A unigene catalogue of 5700 expressed genes in cassava
Plant Mol. Biol. 56 (4), 541-554 (2004)
Contact: Verdier V
Laboratoire genome et developpement des plantes, UMR5096
IRD-CNRS-Universite de Perpignan-CIAT
52 AV Paul Alduy Perpignan 66860 France
Tel: (33) 4 68 66 17 74
Fax: (33) 4 68 66 84 99
Email: Valerie.Verdier@ird.fr
Seq primer: T3.
Location/Qualifiers
source
1..548
/organism="Manihot esculenta"
/mol_type="mRNA"
/cultivar="MPer183"
/db_xref="taxon:3983"
/clone_lib="MPer183 cassava lambda zap"
/note="Organ: Root; Vector: pBluescript SK (+/-); Site_1:
EcoRI; Site_2: XhoI; Stratagene lambda Zap Library"

ORIGIN

Query Match 79.2%; Score 19.8; DB 7; Length 451;
Best Local Similarity 91.3%; Pred. No. 5.5e+02;
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AAAACAATGACCAACTGGCTG 24
|||||
Db 300 AAAACAATGACCAACTGAGCTG 322
|||||

RESULT 20
CK650237
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

Manihot esculenta (cassava)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids I; Malpighiales; Euphorbiales; Crotonoideae;
Manihoteae; Manihot.
1 (bases 1 to 548)
Lopez,C., Jorge,V., Piegue,B., Mba,C., Cortes,D., Restrepo,S.,
Soto,M., Laudie,M., Berger,C., Cooke,R., Delsey,M., Tohme,J. and
Verdier,V.
A unigene catalogue of 5700 expressed genes in cassava
Plant Mol. Biol. 56 (4), 541-554 (2004)
Contact: Verdier V
Laboratoire genome et developpement des plantes, UMR5096
IRD-CNRS-Universite de Perpignan-CIAT
52 AV Paul Alduy Perpignan 66860 France
Tel: (33) 4 68 66 17 74
Fax: (33) 4 68 66 84 99
Email: Valerie.Verdier@ird.fr
Seq primer: T3.
Location/Qualifiers
source
1..548
/organism="Manihot esculenta"
/mol_type="mRNA"
/cultivar="MPer183"
/db_xref="taxon:3983"
/clone_lib="MPer183 cassava lambda zap"
/note="Organ: Root; Vector: pBluescript SK (+/-); Site_1:
EcoRI; Site_2: XhoI; Stratagene lambda Zap Library"

ORIGIN

Query Match 79.2%; Score 19.8; DB 7; Length 548;
Best Local Similarity 91.3%; Pred. No. 5.6e+02;
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AAAACAATGACCAACTGGCTG 24
|||||
Db 435 AAAACAATGACCAACTGAGCTG 457
|||||

RESULT 20
CK650237
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

Manihot esculenta (cassava)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids I; Malpighiales; Euphorbiales; Crotonoideae;
Manihoteae; Manihot.
1 (bases 1 to 548)
Lopez,C., Jorge,V., Piegue,B., Mba,C., Cortes,D., Restrepo,S.,
Soto,M., Laudie,M., Berger,C., Cooke,R., Delsey,M., Tohme,J. and
Verdier,V.
A unigene catalogue of 5700 expressed genes in cassava
Plant Mol. Biol. 56 (4), 541-554 (2004)
Contact: Verdier V
Laboratoire genome et developpement des plantes, UMR5096
IRD-CNRS-Universite de Perpignan-CIAT
52 AV Paul Alduy Perpignan 66860 France
Tel: (33) 4 68 66 17 74
Fax: (33) 4 68 66 84 99
Email: Valerie.Verdier@ird.fr
Seq primer: T3.
Location/Qualifiers
source
1..548
/organism="Manihot esculenta"
/mol_type="mRNA"
/cultivar="MPer183"
/db_xref="taxon:3983"
/clone_lib="MPer183 cassava lambda zap"
/note="Organ: Root; Vector: pBluescript SK (+/-); Site_1:
EcoRI; Site_2: XhoI; Stratagene lambda Zap Library"

ORIGIN

Query Match 79.2%; Score 19.8; DB 7; Length 548;
Best Local Similarity 91.3%; Pred. No. 5.6e+02;
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AAAACAATGACCAACTGGCTG 24
|||||
Db 435 AAAACAATGACCAACTGAGCTG 457
|||||

```

RESULT 21
CF178286/c
LOCUS       CF178286               683 bp    mRNA    linear    EST 28-JUL-2003
DEFINITION   807338 MARC 3PIG Sus scrofa cDNA 5', mRNA sequence.
ACCESSION   CF178286
VERSION     CF178286.1   GI:33290062
KEYWORDS    EST.
SOURCE      Sus scrofa (pig)
ORGANISM    Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae;
            Sus.
REFERENCE   1 (bases 1 to 683)
AUTHORS     Smith,T.P.L., Freking,B.A., Ford,J.J., Vallet,J.L., Fox,J.,
            Wise,I.A., Nonneman,B.J., Wray,J.E. and Keele,J.W.
TITLE       A second set of porcine ESTs from a pooled-tissue normalized
            library
JOURNAL     Unpublished (2003)
COMMENT     Contact: Smith TPL
            USDA, ARS, US Meat Animal Research Center
            FO Box 166, Clay Center, NE 68933-0166, USA
            Tel: 402 762 4366
            Fax: 402 762 4390
            Email: smith@email.marc.usda.gov
            Single pass sequencing. Bases called with phred v0.020425.c and
            trimmed with the aid of the trim_alt option. Vector identified with
            cross_match v0.990329.
            Plate: SRG8006 row: B column: 22
            Seq primer: GTAATACGACTCACTATAGG.
FEATURES             Location/Qualifiers
     source          1..683
                     /organism="Sus scrofa"
                     /mol_type="mRNA"
                     /db_xref="taxon:9823"
                     /tissue_type="pooled"
                     /lab_host="DH10B"
                     /clone_lib="MARC 3PIG"
                     /notes="Vector: pCDNA3.1; Site 1: EcoRI; Site 2: NotI;
                     Library made with RNA pooled from multiple tissues
                     including brain, liver, muscle, placenta/endometrium,
                     ovary, testes, and bone marrow."
ORIGIN
Query Match       79.2%;   Score 19.8;   DB 6;   Length 683;
Best Local Similarity 91.3%;   Pred. No. 5.8e+02;
Matches 21;   Conservative 0;   Mismatches 2;   Indels 0;   Gaps 0;

QY       1  AAAAAACATGACCAACTGGGCT 23
          |||||  |||||  |||||  |||||  |||||
Db       323 AAAAAACAGGACCAACTGGGCT 301

RESULT 22
CK011470/c
LOCUS       CK011470               422 bp    mRNA    linear    EST 04-FEB-2005
DEFINITION   29146rsicef_7286.y1 Oryza sativa cv. PA64s panicle sterile cDNA
            library Oryza sativa (indica cultivar-group) cDNA 5', mRNA
            sequence.
ACCESSION   CK011470
VERSION     CK011470.1   GI:58600942
KEYWORDS    EST.
SOURCE      Oryza sativa (indica cultivar-group)
ORGANISM    Oryza sativa (indica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE   1 (bases 1 to 422)
AUTHORS     Yu,J., Wang,J., Lin,W., Li,S., Li,H., Zhou,J., Ni,P., Dong,W.,
            Hu,S., Zeng,C., Zhang,J., Zhang,Y., Li,R., Xu,Z., Li,S., Li,X.,
            Zheng,H., Cong,L., Lin,L., Yin,J., Geng,J., Li,G., Shi,J., Liu,J.,
            Lv,H., Li,J., Wang,J., Deng,Y., Ran,L., Shi,X., Wang,X., Wu,Q.,
            Li,C., Ren,X., Wang,J., Wang,X., Li,D., Liu,D., Zhang,X., Ji,Z.,
            Zhao,W., Sun,Y., Zhang,Z., Bao,J., Han,Y., Dong,L., Ji,J., Chen,P.,
            Wu,S. and Liu,J.
TITLE       The Genomes of Oryza sativa: A History of Duplications
JOURNAL     PLoS Biol. 3 (2), e38 (2005)
COMMENT     Contact: Yan Zhou
            Bioinformatics Department
            Hangzhou Genomics Institute
            No.51 Zhijiang Road, Hangzhou 310008, China
            Tel: 86-571-56805886
            Fax: 86-571-56805884
            Email: zhouyan@genomics.org.cn
            Seq primer: M13 Forward
            High quality sequence stop: 422
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1..422
                     /organism="Oryza sativa (indica cultivar-group)"
                     /mol_type="mRNA"
                     /cultivar="PA64s"
                     /db_xref="taxon:39946"
                     /tissue_type="panicle"
                     /cell_type="sterile"
                     /dev_stage="heading/flowering"
                     /clone_lib="Oryza sativa cv. PA64s panicle sterile cDNA
                     library"
ORIGIN
Query Match       77.6%;   Score 19.4;   DB 7;   Length 422;
Best Local Similarity 95.2%;   Pred. No. 8.3e+02;
Matches 20;   Conservative 0;   Mismatches 1;   Indels 0;   Gaps 0;

QY       3  AAAACAATGACCAACTGGGCT 23
          |||||  |||||  |||||  |||||  |||||
Db       123 AAAATAATGACCAACTGGGCT 103

RESULT 23
CK063363/c
LOCUS       CK063363               662 bp    mRNA    linear    EST 05-FEB-2005
DEFINITION   73210rsicem_6604.y1 Oryza sativa cv. PA64s panicle sterile cDNA
            library Oryza sativa (indica cultivar-group) cDNA 5', mRNA
            sequence.
ACCESSION   CK063363
VERSION     CK063363.1   GI:58674676
KEYWORDS    EST.
SOURCE      Oryza sativa (indica cultivar-group)
ORGANISM    Oryza sativa (indica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE   1 (bases 1 to 662)
AUTHORS     Yu,J., Wang,J., Lin,W., Li,S., Li,H., Zhou,J., Ni,P., Dong,W.,
            Hu,S., Zeng,C., Zhang,J., Zhang,Y., Li,R., Xu,Z., Li,S., Li,X.,
            Zheng,H., Cong,L., Lin,L., Yin,J., Geng,J., Li,G., Shi,J., Liu,J.,
            Lv,H., Li,J., Wang,J., Deng,Y., Ran,L., Shi,X., Wang,X., Wu,Q.,
            Li,C., Ren,X., Wang,J., Wang,X., Li,D., Liu,D., Zhang,X., Ji,Z.,
            Zhao,W., Sun,Y., Zhang,Z., Bao,J., Han,Y., Dong,L., Ji,J., Chen,P.,
            Wu,S. and Liu,J.
TITLE       The Genomes of Oryza sativa: A History of Duplications
JOURNAL     PLoS Biol. 3 (2), e38 (2005)
COMMENT     Contact: Yan Zhou
            Bioinformatics Department
            Hangzhou Genomics Institute
            No.51 Zhijiang Road, Hangzhou 310008, China
            Tel: 86-571-56805886
            Fax: 86-571-56805884
            Email: zhouyan@genomics.org.cn
            Seq primer: M13 Forward
            High quality sequence stop: 662
            POLYA=No.
FEATURES             Location/Qualifiers

```

Zhao,W., Sun,Y., Zhang,Z., Bao,J., Han,Y., Dong,L., Ji,J., Chen,P.,
 Wu,S. and Liu,J.
 The Genomes of Oryza sativa: A History of Duplications
 PLoS Biol. 3 (2), e38 (2005)
 15685292
 Contact: Yan Zhou
 Bioinformatics Department
 Hangzhou Genomics Institute
 No.51 Zhijiang Road, Hangzhou 310008, China
 Tel: 86-571-56805886
 Fax: 86-571-56805884
 Email: zhouyan@genomics.org.cn
 Seq primer: M13 Forward
 High quality sequence stop: 422
 POLYA=No.

FEATURES
 source

1..422
 /organism="Oryza sativa (indica cultivar-group)"
 /mol_type="mRNA"
 /cultivar="PA64s"
 /db_xref="taxon:39946"
 /tissue_type="panicle"
 /cell_type="sterile"
 /dev_stage="heading/flowering"
 /clone_lib="Oryza sativa cv. PA64s panicle sterile cDNA
 library"

ORIGIN

Query Match 77.6%; Score 19.4; DB 7; Length 422;
 Best Local Similarity 95.2%; Pred. No. 8.3e+02;
 Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 AAAACAATGACCAACTGGGCT 23

||||| ||||| ||||| ||||| |||||
 Db 123 AAAATAATGACCAACTGGGCT 103

RESULT 23

CK063363/c
 LOCUS
 DEFINITION

73210rsicem_6604.y1 Oryza sativa cv. PA64s panicle sterile cDNA
 library Oryza sativa (indica cultivar-group) cDNA 5', mRNA
 sequence.

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

Oryza sativa (indica cultivar-group)
 Oryza sativa (indica cultivar-group)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE
 AUTHORS

1 (bases 1 to 662)
 Yu,J., Wang,J., Lin,W., Li,S., Li,H., Zhou,J., Ni,P., Dong,W.,
 Hu,S., Zeng,C., Zhang,J., Zhang,Y., Li,R., Xu,Z., Li,S., Li,X.,
 Zheng,H., Cong,L., Lin,L., Yin,J., Geng,J., Li,G., Shi,J., Liu,J.,
 Lv,H., Li,J., Wang,J., Deng,Y., Ran,L., Shi,X., Wang,X., Wu,Q.,
 Li,C., Ren,X., Wang,J., Wang,X., Li,D., Liu,D., Zhang,X., Ji,Z.,
 Zhao,W., Sun,Y., Zhang,Z., Bao,J., Han,Y., Dong,L., Ji,J., Chen,P.,
 Wu,S. and Liu,J.

TITLE
 JOURNAL
 PUBMED
 COMMENT

The Genomes of Oryza sativa: A History of Duplications
 PLoS Biol. 3 (2), e38 (2005)
 15685292

Contact: Yan Zhou

Bioinformatics Department
 Hangzhou Genomics Institute
 No.51 Zhijiang Road, Hangzhou 310008, China
 Tel: 86-571-56805886
 Fax: 86-571-56805884

Email: zhouyan@genomics.org.cn

Seq primer: M13 Forward

High quality sequence stop: 662

POLYA=No.

FEATURES

Location/Qualifiers

```

source
1. .662
/organism="Oryza sativa (indica cultivar-group)"
/mol_type="mRNA"
/db_xref="taxon:39946"
/tissue_type="panicle"
/cell_type="fertile"
/dev_stage="heading/flowering"
/clone_lib="Oryza sativa cv. PA64s panicle sterile cDNA
library"

ORIGIN
Query Match 77.6%; Score 19.4; DB 7; Length 662;
Best Local Similarity 95.2%; Pred. No. 8.7e+02;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 AAAACAATGACCAACTGGGCT 23
|||||
Db 538 AAAATAATGACCAACTGGGCT 518

RESULT 24
CK009716/c
LOCUS
DEFINITION
33657ricef 6041.y1 Oryza sativa cv. PA64s panicle sterile cDNA
library Oryza sativa (indica cultivar-group) cDNA 5', mRNA
sequence.
ACCESSION
CK009716
VERSION
CK009716.1 GI:58599188
KEYWORDS
EST.
SOURCE
Oryza sativa (indica cultivar-group)
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE
1 (bases 1 to 678)
Yu,J., Wang,J., Lin,W., Li,S., Li,H., Zhou,J., Ni,P., Dong,W.,
Hu,S., Zeng,C., Zhang,J., Zhang,Y., Li,R., Xu,Z., Li,S., Li,X.,
Zheng,H., Cong,L., Lin,L., Yin,J., Geng,J., Li,G., Shi,J., Liu,J.,
Lv,H., Li,J., Wang,J., Deng,Y., Ran,L., Shi,X., Wang,X., Wu,Q.,
Li,C., Ren,X., Wang,J., Wang,X., Li,D., Liu,D., Zhang,X., Ji,Z.,
Zhao,W., Sun,Y., Zhang,Z., Bao,J., Han,Y., Dong,L., Ji,J., Chen,P.,
Wu,S. and Liu,J.
The Genomes of Oryza sativa: A History of Duplications
PLOS Biol. 3 (2), e38 (2005)
15685292
Contact: Yan Zhou
Bioinformatics Department
Hangzhou Genomics Institute
No.51 Zhijiang Road, Hangzhou 310008, China
Tel: 86-571-56805886
Fax: 86-571-56805884
Email: zhouyan@genomics.org.cn
Seq primer: M13 Forward
High quality sequence stop: 678
POLYA=No.

FEATURES
source
1. .678
/organism="Oryza sativa (indica cultivar-group)"
/mol_type="mRNA"
/db_xref="taxon:39946"
/db_xref="taxon:39946"
/tissue_type="panicle"
/cell_type="sterile"
/dev_stage="heading/flowering"
/clone_lib="Oryza sativa cv. PA64s panicle sterile cDNA
library"

ORIGIN
Query Match 77.6%; Score 19.4; DB 7; Length 678;
Best Local Similarity 95.2%; Pred. No. 8.7e+02;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

source
1. .662
/organism="Oryza sativa (indica cultivar-group)"
/mol_type="mRNA"
/db_xref="taxon:39946"
/tissue_type="panicle"
/cell_type="fertile"
/dev_stage="heading/flowering"
/clone_lib="Oryza sativa cv. PA64s panicle sterile cDNA
library"

ORIGIN
Query Match 77.6%; Score 19.4; DB 9; Length 704;
Best Local Similarity 95.2%; Pred. No. 8.8e+02;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 5 AACCAATGACCAACTGGGCTGT 25
|||||
Db 533 AACCAATGACCAACTGGTCTGT 553

RESULT 26
CB677845
LOCUS
DEFINITION
OSJUNE15D19.r OSJUNE Oryza sativa (japonica cultivar-group) cDNA
clone OSJUNE15D19 3', mRNA sequence.
ACCESSION
CB677845
VERSION
CB677845.1 GI:29681570
KEYWORDS
EST.
SOURCE
Oryza sativa (japonica cultivar-group)
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE
1 (bases 1 to 812)
Jantassuriyarat,C., Gowda,M., Haller,K., Hatfield,J., Lu,G.,
Stahlberg,E., Zhou,B., Li,H., Kim,H., Yu,Y., Dean,R.A., Wing,R.A.,
Soderlund,C. and Wang,G.L.

```

```

Qy 3 AAAACAATGACCAACTGGGCT 23
|||||
Db 149 AAAATAATGACCAACTGGGCT 129

RESULT 25
BZ017812
LOCUS
DEFINITION
oe66607.b1 B.oleracea002 Brassica oleracea genomic, genomic survey
sequence.
ACCESSION
BZ017812
VERSION
BZ017812.1 GI:23576701
KEYWORDS
GSS.
SOURCE
Brassica oleracea
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE
1 (bases 1 to 704)
Delehaunty,K., Fewell,G., Fulton,L., McCombie,W.R., Miner,T.,
Nash,W., Rabinowicz,P.D. and Wilson,R.K.
Whole genome shotgun reads from Brassica oleracea
Unpublished (2002)
Contact: Richard K. Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@watson.wustl.edu
Plate: oek66 row: d column: 07
Seq primer: -21uppot forward
Class: shotgun
High quality sequence start: 16
High quality sequence stop: 551.
Location/Qualifiers
1. .704
/organism="Brassica oleracea"
/mol_type="genomic DNA"
/db_xref="taxon:3712"
/clone_lib="B.oleracea002"
/notes="vector: potW13; Whole genome shotgun library from
flowering buds. DNA was purified from a crude nuclear
prep using Brassica oleracea TO1000DH3 buds provided by
Thomas Osborn at the University of Wisconsin. Genomic
DNA was provided by Pablo Rabinowicz (CSHL) and the
shotgun library prepared at Washington University Genome
Sequencing Center."

ORIGIN
Query Match 77.6%; Score 19.4; DB 9; Length 704;
Best Local Similarity 95.2%; Pred. No. 8.8e+02;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 5 AACCAATGACCAACTGGGCTGT 25
|||||
Db 533 AACCAATGACCAACTGGTCTGT 553

RESULT 26
CB677845
LOCUS
DEFINITION
OSJUNE15D19.r OSJUNE Oryza sativa (japonica cultivar-group) cDNA
clone OSJUNE15D19 3', mRNA sequence.
ACCESSION
CB677845
VERSION
CB677845.1 GI:29681570
KEYWORDS
EST.
SOURCE
Oryza sativa (japonica cultivar-group)
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE
1 (bases 1 to 812)
Jantassuriyarat,C., Gowda,M., Haller,K., Hatfield,J., Lu,G.,
Stahlberg,E., Zhou,B., Li,H., Kim,H., Yu,Y., Dean,R.A., Wing,R.A.,
Soderlund,C. and Wang,G.L.

```

TITLE Large-scale identification of expressed sequence tags involved in rice and rice blast fungus interaction
JOURNAL Plant Physiol. 138 (1), 105-115 (2005)
PUBMED 1588683
COMMENT Contact: Rod Wing
 Arizona Genomics Institute
 University of Arizona
 Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ 85721-0088, USA
 Tel: 520 626 3967
 Fax: 520 621 3288
 Email: rwing@genome.arizona.edu

PCR Primers
 FORWARD: gta aac cga cgg cca gtg
 BACKWARD: gga aac agc tat gac cat g
 Plate: 15 row: D column: 19
 Seq primer: gga aac agc tat gac cat g.

Location/Qualifiers
 1. .812
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="mRNA"
 /cultivar="Nipponbare"
 /db_xref="taxon:39947"
 /clone="OSJNEe15D19"
 /tissue_type="Leaf"
 /dev_stage="3 week"
 /lab_host="DH10B"
 /clone_lib="OSJNEe"
 /notes="Vector: pBluescript II KS +; Site 1: EcoRI; Site 2: XhoI; 24 hrs after inoculation with Rice Blast (70-15)"

FEATURES

source
 Query Match 77.6%; Score 19.4; DB 6; Length 812;
 Best Local Similarity 95.2%; Pred. No. 8.9e+02;
 Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 AAAACAATGACCAACTGGGCT 23
 |||||
 Db 775 AAAATATGACCAACTGGGCT 795

RESULT 27

CB6777843
LOCUS OSJNEe15D18 r OSJNEe Oryza sativa (japonica cultivar-group) cDNA
DEFINITION clone OSJNEe15D18 3', mRNA sequence.
ACCESSION CB6777843.1 GI:29681568
VERSION EST.
KEYWORDS Oryza sativa (japonica cultivar-group)
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE

AUTHORS Jantasuriyarat, C., Gowda, M., Haller, K., Hatfield, J., Lu, G., Stahlberg, E., Zhou, B., Li, H., Kim, H., Yu, Y., Dean, R.A., Wing, R.A., Soderlund, C. and Wang, G.L.
TITLE Large-scale identification of expressed sequence tags involved in rice and rice blast fungus interaction
JOURNAL Plant Physiol. 138 (1), 105-115 (2005)
PUBMED 1588683
COMMENT Contact: Rod Wing
 Arizona Genomics Institute
 University of Arizona
 Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ 85721-0088, USA
 Tel: 520 626 3967
 Fax: 520 621 3288
 Email: rwing@genome.arizona.edu
 PCR Primers
 FORWARD: gta aac cga cgg cca gtg
 BACKWARD: gga aac agc tat gac cat g

Plate: 15 row: D column: 18
 Seq primer: gga aac agc tat gac cat g.
 Location/Qualifiers
 1. .911
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="mRNA"
 /cultivar="Nipponbare"
 /db_xref="taxon:39947"
 /clone="OSJNEe15D18"
 /tissue_type="Leaf"
 /dev_stage="3 week"
 /lab_host="DH10B"
 /clone_lib="OSJNEe"
 /notes="Vector: pBluescript II KS +; Site 1: EcoRI; Site 2: XhoI; 24 hrs after inoculation with Rice Blast (70-15)"

FEATURES

source
 Query Match 77.6%; Score 19.4; DB 6; Length 911;
 Best Local Similarity 95.2%; Pred. No. 9.1e+02;
 Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 AAAACAATGACCAACTGGGCT 23
 |||||
 Db 828 AAAATATGACCAACTGGGCT 848

ORIGIN

Query Match 77.6%; Score 19.4; DB 6; Length 911;
 Best Local Similarity 95.2%; Pred. No. 9.1e+02;
 Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 AAAACAATGACCAACTGGGCT 23
 |||||
 Db 828 AAAATATGACCAACTGGGCT 848

RESULT 28

CF330489
LOCUS NACL--06-D15.g1 Rice callus plasmid cDNA library (NACL) Oryza sativa (japonica cultivar-group) cDNA clone NACL--06-D15, mRNA sequence.
ACCESSION CF330489.1 GI:33809212
VERSION EST.
KEYWORDS Oryza sativa (japonica cultivar-group)
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE

AUTHORS Kim, J.S., Jun, K.M., Cheong, P.J., Kim, M.J., Lee, T.H., Shin, Y.C., Song, S.I., Kim, J.K., Kim, Y.-K. and Nahm, B.H.
TITLE Large-scale Sequencing Analysis of Rice ESTs
JOURNAL Unpublished (2003)
COMMENT Contact: Nahm B.H.
 Genomics and Genetics Institute, GreenGene Biotech Inc.; Division of Bioscience and Bioinformatics, Myongji University
 Yongin, Gyeonggi, Korea
 Tel: 82 31 330 6193
 Fax: 82 31 321 6355
 Email: bhnahm@gbio.com, bhnahm@bio.myongji.ac.kr.

FEATURES

source
 Query Match 76.8%; Score 19.2; DB 6; Length 163;
 Best Local Similarity 87.5%; Pred. No. 9.1e+02;
 Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAACAATGACCAACTGGGCT 25
 |||||

ORIGIN

Query Match 76.8%; Score 19.2; DB 6; Length 163;
 Best Local Similarity 87.5%; Pred. No. 9.1e+02;
 Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAACAATGACCAACTGGGCT 25
 |||||

```

Db      69 AAAAAAGTGTGACCAACTGGGCTGT 92
|||||  |||||||  |||||||  |||||||  |||||||
COMMENT Other GSSs: CH240_182B22.TJ
Contact: Caetano AR
Department of Biotechnology
Embrapa Recursos Geneticos e Biotecnologia
Parque Estacao Biologica, Final Av. W/5 Norte, Brasilia-DF C.P.
02372, 70770-900 Brasil
Tel: 55 61 448 4778
Fax: 55 61 340 1658
Email: acetanoc@embrapa.br
Clones are derived from the bovine BAC library CHORI-240
(http://www.chori.org/bacpac/bovine240.htm).
Bases shown have phred quality value equal to or higher than 20.
Bases with quality value below 20 were masked with 'N'.
For BAC library availability, please contact Pieter de Jong
(pdejong@mail.choi.org).
Clones may be purchased from BACPAC Resources
(http://www.chori.org/bacpac/ordering_information.htm).
This work was undertaken as part of the International Bovine BAC
Mapping Consortium (IBBMC) by Embrapa Recursos Geneticos e
Biotecnologia with financing from Conselho Nacional de
Desenvolvimento Cientifico e Tecnologico (CNPq), Brazil
Plate: 182 row: B column: 22
Seq primer: T7
Class: BAC ends
High quality sequence stop: 471.
Location/Qualifiers
1..471
/organism="Bos taurus"
/mol_type="genomic DNA"
/strain="breed: Hereford"
/db_xref="taxon:9913"
/clone="CH240_182B22"
/sex="Male"
/cell_type="Blood"
/clone_lib="CHORI-240"
/note="Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI;
Hereford bull L1 Domino 99375; CHORI-240 Bovine BAC
library (Male) produced by Pieter de Jong"

FEATURES             source
ORIGIN
Query Match          76.8%; Score 19.2; DB 10; Length 471;
Best Local Similarity 87.5%; Pred. No. 1e+03;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAACAATGACCAACTGGGCTG 24
|||||  |||||||  |||||||  |||||||  |||||||
Db 334 AAAAAACAATGATCAAGTGGACTG 311
|||||  |||||||  |||||||  |||||||  |||||||

RESULT 31
AZ307205/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Mus musculus (house mouse)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 472)
Dunn, D., Aoyagi, A., Barber, M., Becorn, T., Duval, B., Hamil, C.,
Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T.,
Reilly, M., Rose, R., Rose, R., Stokes, R., Tingey, A., von
Niederhausern, A. and Wright, D., Weiss, R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah

COMMENT

Db 95 AAAAAAGTGTGACCAACTGGGCTGT 72
|||||  |||||||  |||||||  |||||||  |||||||
COMMENT Other GSSs: CH240_182B22.TJ
Contact: Caetano AR
Department of Biotechnology
Embrapa Recursos Geneticos e Biotecnologia
Parque Estacao Biologica, Final Av. W/5 Norte, Brasilia-DF C.P.
02372, 70770-900 Brasil
Tel: 55 61 448 4778
Fax: 55 61 340 1658
Email: acetanoc@embrapa.br
Clones are derived from the bovine BAC library CHORI-240
(http://www.chori.org/bacpac/bovine240.htm).
Bases shown have phred quality value equal to or higher than 20.
Bases with quality value below 20 were masked with 'N'.
For BAC library availability, please contact Pieter de Jong
(pdejong@mail.choi.org).
Clones may be purchased from BACPAC Resources
(http://www.chori.org/bacpac/ordering_information.htm).
This work was undertaken as part of the International Bovine BAC
Mapping Consortium (IBBMC) by Embrapa Recursos Geneticos e
Biotecnologia with financing from Conselho Nacional de
Desenvolvimento Cientifico e Tecnologico (CNPq), Brazil
Plate: 182 row: B column: 22
Seq primer: T7
Class: BAC ends
High quality sequence stop: 471.
Location/Qualifiers
1..471
/organism="Bos taurus"
/mol_type="genomic DNA"
/strain="breed: Hereford"
/db_xref="taxon:9913"
/clone="CH240_182B22"
/sex="Male"
/cell_type="Blood"
/clone_lib="CHORI-240"
/note="Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI;
Hereford bull L1 Domino 99375; CHORI-240 Bovine BAC
library (Male) produced by Pieter de Jong"

FEATURES             source
ORIGIN
Query Match          76.8%; Score 19.2; DB 6; Length 192;
Best Local Similarity 87.5%; Pred. No. 9.2e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAAACAATGACCAACTGGGCTGT 25
|||||  |||||||  |||||||  |||||||  |||||||
Db 95 AAAAAAGTGTGACCAACTGGGCTGT 72
|||||  |||||||  |||||||  |||||||  |||||||

RESULT 30
CL605612/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Bos taurus (cow)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
Pecora; Bovidae; Bovinae; Bos.
1 (bases 1 to 471)
Costa, J.N., Mota, M. and Caetano, A.R.
Brazil's Contribution to End-Sequencing the Bovine BAC Library
CHORI-240
Unpublished (2003)

JOURNAL

```


/clone_lib="Rice panicle at flowering stage"
/note="Organ: panicle; Rice cDNA from panicle at flowering stage"

ORIGIN

Query Match 76.8%; Score 19.2; DB 1; Length 709;
Best Local Similarity 87.5%; Pred. No. 1.1e+03;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAAAATGACCAACTGGGCTGT 25

Db 631 AAAAAGTGTGACCAACTGGGCTGT 654

RESULT 34

LOCUS CX248961/c

DEFINITION 1301028 NCCWA 02RT Oncorhynchus mykiss cDNA 3', mRNA sequence.

ACCESSION CX248961

VERSION CX248961.1 GI:60365483

SOURCE EST.

ORGANISM Oncorhynchus mykiss (rainbow trout)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Actinopterygii; Neopterygii; Teleostei; Euteleostei;

Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.

1 (bases 1 to 719)

Rexroad, C.E., Goupil, A.-S., Guiguen, Y. and Yao, J.

02RT IUS, NCCWA/WVU EST Project, Phase II, in collaboration with

INRA

Unpublished (2004)

Contact: Rexroad CE

USDA, ARS, National Center for Cool and Cold Water Aquaculture

11876 Leetown Road, Kearneysville, WV 25430, USA

Tel: 304 724 8340 x2129

Fax: 304 725 0351

Email: crexroad@nccwa.ars.usda.gov

Single pass sequencing. Bases called with phred v0.020425.c and

trimmed with the aid of the trim_alt option. Vector identified with

cross_match v0.990329.

Plate: 106 row: F column: 4

Seq primer: GTAATACGACTCACTATAGG.

FEATURES

source

Location/Qualifiers

1..719

/organism="Oncorhynchus mykiss"

/mol_type="mRNA"

/db_xref="taxon:8022"

/tissue_type="pooled"

/lab_host="DH10B"

/clone_lib="NCCWA 02RT"

/note="Vector: pCMV Sport6.0; This library was created by

A.-S. Goupil and Y. Guiguen who subtracted the NCCWA 1RT

library from the INRA multi-tissue library."

ORIGIN

Query Match 76.8%; Score 19.2; DB 8; Length 719;
Best Local Similarity 87.5%; Pred. No. 1.1e+03;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAAATGACCAACTGGGCTG 24

Db 287 AAAAAAATGACCAACTGGGCTG 264

RESULT 35

LOCUS CR058845/c

DEFINITION 779 bp DNA linear GSS 05-JUL-2004

Reverse strand read from insert in 3'HPRT insertion targeting and

chromosome engineering clone MHP270a06, genomic survey sequence.

ACCESSION CR058845

VERSION CR058845.1 GI:49792317

KEYWORDS GSS; genome survey sequence; MICER.

SOURCE Mus musculus (house mouse)

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

Sciurognathi; Muroidae; Murinae; Mus.

1 (bases 1 to 779)

Adams, D.J., Biggs, P.J., Cox, A.V., Davies, R.M., van der Weyden, L.,

Jonkers, J., Smith, J., Plumb, R.W., Taylor, R.G., Nishijima, I., Yu, Y.,

Rogers, J. and Bradley, A.

Direct Submission

Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,

CB10 1SA, UK. http://www.sanger.ac.uk/MICER

Location/Qualifiers

1..779

/organism="Mus musculus"

/mol_type="genomic DNA"

/db_xref="taxon:10090"

/clone="MHP270a06"

/clone_lib="MHPP"

ORIGIN

Query Match 76.8%; Score 19.2; DB 11; Length 779;
Best Local Similarity 87.5%; Pred. No. 1.1e+03;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAAAATGACCAACTGGGCTGT 25

Db 72 AAAAAAATGACCAACTGGGCTGT 49

RESULT 36

LOCUS CR021797/c

DEFINITION 792 bp DNA linear GSS 05-JUL-2004

Reverse strand read from insert in 3'HPRT insertion targeting and

chromosome engineering clone MHP270a06, genomic survey sequence.

ACCESSION CR021797

VERSION CR021797.1 GI:49754852

KEYWORDS GSS; genome survey sequence; MICER.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

Sciurognathi; Muroidae; Murinae; Mus.

1 (bases 1 to 792)

Adams, D.J., Biggs, P.J., Cox, A.V., Davies, R.M., van der Weyden, L.,

Jonkers, J., Smith, J., Plumb, R.W., Taylor, R.G., Nishijima, I., Yu, Y.,

Rogers, J. and Bradley, A.

Direct Submission

Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,

CB10 1SA, UK. http://www.sanger.ac.uk/MICER

Location/Qualifiers

1..792

/organism="Mus musculus"

/mol_type="genomic DNA"

/db_xref="taxon:10090"

/clone="MHP270a06"

/clone_lib="MHPP"

ORIGIN

Query Match 76.8%; Score 19.2; DB 11; Length 792;
Best Local Similarity 87.5%; Pred. No. 1.1e+03;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAAAATGACCAACTGGGCTGT 25

Db 72 AAAAAAATGACCAACTGGGCTGT 49

RESULT 37

LOCUS CR090689/c

DEFINITION 801 bp DNA linear GSS 05-JUL-2004

Reverse strand read from insert in 3'HPRT insertion targeting and

chromosome engineering clone MHP270a06, genomic survey sequence.

ACCESSION CR090689

```
VERSION CR090689.1 GI:49824440
KEYWORDS GSS: genome survey sequence; MICER.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 801)
AUTHORS Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L.,
Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,
Rogers,J. and Bradley,A.
TITLE Direct Submission
JOURNAL Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. http://www.sanger.ac.uk/MICER
FEATURES
source
1..801
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="MHPP304G17"
/clone_lib="MHPP"
ORIGIN
Query Match 76.8%; Score 19.2; DB 11; Length 801;
Best Local Similarity 87.5%; Pred. No. 1.1e+03;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 2 AAAAAACAATGACCAACTGGGCTGT 25
|||||
Db 72 AAAAAACAATGACCAAGCGGTTGT 49
|||||
RESULT 38
AG479423 808 bp DNA linear GSS 22-DEC-2004
LOCUS Mus musculus molossinus DNA, clone:MSMg01-373G01.TJ, genomic survey
DEFINITION sequence.
ACCESSION AG479423
VERSION AG479423.1 GI:48186653
KEYWORDS GSS.
SOURCE Mus musculus molossinus (Japanese wild mouse)
ORGANISM Mus musculus molossinus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Abe,K., Noguchi,H., Tagawa,K., Yuzuriha,M., Toyoda,A., Kojima,T.,
Ezawa,K., Saitou,N., Hattori,M., Sakaki,Y., Moriawaki,K. and
Shiroishi,T.
TITLE Contribution of Asian mouse subspecies Mus musculus molossinus to
genomic constitution of strain C57BL/6J, as defined by BAC-end
sequence-SNP analysis
JOURNAL Genome Res. 14 (12), 2439-2447 (2004)
PUBMED 15574823
REFERENCE 2 (bases 1 to 808)
AUTHORS Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
TITLE Direct Submission
JOURNAL Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa, 230-0045, Japan
(E-mail:hattori@gsc.riken.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT Clones are derived from the mouse BAC library MSMg01. For BAC
library availability, please contact Kuniya Abe (abe@rtc.riken.jp).
Taikuba Institute, Bio Resource Center,
The Institute of Physical and Chemical Research (RIKEN) 3-1-1
Koyadai, Tsukuba, 305-0074 Japan
phone: 81-298-36-9189, fax: 81-298-36-9199
e-mail: abe@rtc.riken.jp
PRIMERS
Sequencing : TJ
LIBRARY
Vector : pBACe3.6

R.Site 1 : EcoRI
R.Site 2 : EcoRI.
Location/Qualifiers
1..808
/organism="Mus musculus molossinus"
/mol_type="genomic DNA"
/sub_species="molossinus"
/db_xref="taxon:57486"
/clone="MSMg01-373G01.TJ"
/sex="male"
/tissue_type="mixture of kidney and spleen"
/clone_lib="MSMg01 Mouse Male BAC Library"
ORIGIN
Query Match 76.8%; Score 19.2; DB 10; Length 808;
Best Local Similarity 87.5%; Pred. No. 1.1e+03;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 AAAAAACAATGACCAACTGGGCTG 24
|||||
Db 528 AAAAAAATATCACCACCTGGGCTG 505
|||||
RESULT 39
AU252333 811 bp mRNA linear EST 26-JAN-2005
LOCUS AU252333 salt-stressed barley root cDNA Hordeum vulgare subsp.
DEFINITION vulgare cDNA clone BR-T08 similar to HvLRR2, mRNA sequence.
ACCESSION AU252333
VERSION AU252333.1 GI:19494120
KEYWORDS EST.
SOURCE Hordeum vulgare subsp. vulgare
ORGANISM Hordeum vulgare subsp. vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poaceae; Triticeae; Hordeum.
REFERENCE 1 (bases 1 to 811)
AUTHORS Ueda,A., Shi,W., Nakamura,T. and Takabe,T.
TITLE Analysis of salt-inducible genes in barley roots by differential
display
JOURNAL J. Plant Res. 115 (1118), 119-130 (2002)
PUBMED 12884135
COMMENT Contact: Akihiro Ueda
Lab. of Biosphere Symbioses
Graduate School of Bioagricultural Sciences, Nagoya University
Furo-cho, Chikusa, Nagoya, Aichi 464-8601, Japan
Tel: 81-52-789-4044(ex.5209)
Fax: 81-52-789-5209
Email: akihiro@nuagri1.agr.nagoya-u.ac.jp.
FEATURES
source
1..811
/organism="Hordeum vulgare subsp. vulgare"
/mol_type="mRNA"
/cultivar="Haruna-ni-jo"
/sub_species="vulgare"
/db_xref="taxon:112509"
/clone="BR-T08"
/tissue_type="root"
/dev_stage="early vegetative stage"
/clone_lib="salt-stressed barley root cDNA"
ORIGIN
Query Match 76.8%; Score 19.2; DB 1; Length 811;
Best Local Similarity 87.8%; Pred. No. 1.1e+03;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 AAAAAACAATGACCAACTGGGCTG 24
|||||
Db 249 AAGATAAATGACCAACTGGGCTG 272
|||||
RESULT 40
CR101220/c
```



```

EST.
SOURCE  Homo sapiens (human)
ORGANISM

REFERENCE
AUTHORS  Li,W.B., Gruber,C., Jesse,J. and Polayes,D.
TITLE    Full-length cDNA libraries and normalization
JOURNAL  Unpublished (2001)
COMMENT  On May 15, 2003 this sequence version replaced gi:30771985.

Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Crémieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
This sequence belongs to sequence cluster 3595.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0DF023AC030P1kc=3595.r.

FEATURES
source
Location/Qualifiers
1..990
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DF023YE05"
/tissue_type="FETAL BRAIN"
/dev_stage="fetal"
/clone_lib="Homo sapiens FETAL BRAIN"
/notes="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-strand cDNA was digested with Not I and
cloned into the Not I and EcoRV sites of the pCMVSPORT 6
vector. Library was not normalized."

ORIGIN

Query Match 76.0%; Score 19; DB 5; Length 990;
Best Local Similarity 76.0%; Pred. No. 1.4e+03;
Matches 19; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAAATGACCAACTGGCGTGT 25
|||||:::|||||
747 AAAAAAAAWMGCACTGGCGGT 771

RESULT 44
LOCUS  CNS00LH5 1101 bp DNA linear GSS 03-JUN-1999
DEFINITION Drosophila melanogaster genome survey sequence 17 end of BAC:
BACR30106 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION AL068211
VERSION AL068211.1 GI:4958428
KEYWORDS GSS.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 1101)
/dev_stage="fetal"
/clone_lib="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/clone="BACR30106"
/clone_lib="RPCI-98"
/notes="end : 17"

FEATURES
source
Location/Qualifiers
1..1101
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR30106"
/clone_lib="RPCI-98"
/notes="end : 17"

ORIGIN

Query Match 76.0%; Score 19; DB 10; Length 1101;
Best Local Similarity 76.0%; Pred. No. 1.4e+03;
Matches 19; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAAATGACCAACTGGCGTGT 25
|||||:::|||||
450 AAAAAAAACCAACTGGCGTGT 474

RESULT 45
LOCUS  BW779425/c 333 bp mRNA linear EST 10-AUG-2005
DEFINITION BW779425 Amphioxus Branchiostoma floridae unpublished cDNA library,
gastrula whole animal Branchiostoma floridae cDNA clone bbg0055a01
5', mRNA sequence.
ACCESSION BW779425
VERSION BW779425.1 GI:66381919
KEYWORDS EST.
SOURCE Branchiostoma floridae (Florida lancelet)
ORGANISM Branchiostoma floridae
Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;
Branchiostoma.
1 (bases 1 to 333)
/notes="1 to 333"
AUTHORS Yu,J., Holland,L.Z., Shin-i,T., Kohara,Y., Satou,Y. and Satoh,N.
TITLE Expressed genes in Branchiostoma floridae
JOURNAL Unpublished (2005)
COMMENT Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshin@genes.nig.ac.jp
If you want to have a cDNA clone for this EST or if you have any
questions, please send an e-mail to Nori Satoh
(satoh@ascidian.zool.kyoto-u.ac.jp) and its cc to Yutaka Satou
(yutaka@ascidian.zool.kyoto-u.ac.jp).

FEATURES
source
Location/Qualifiers
1..333
/organism="Branchiostoma floridae"
/mol_type="mRNA"
/db_xref="taxon:7739"
/clone="bbg0055a01"
/tissue_type="whole animal"
/dev_stage="gastrula"
/clone_lib="Amphioxus Branchiostoma floridae unpublished
cDNA library, gastrula whole animal"

ORIGIN

Query Match 75.2%; Score 18.8; DB 5; Length 333;
Best Local Similarity 87.0%; Pred. No. 1.5e+03;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

melanogaster BAC library was prepared by Kazutoyo Oosagawa and Aaron Mammeter in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw 9p, the same strain used for the BDGP's p1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES

source

```

Location/Qualifiers
1..1101
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR30106"
/clone_lib="RPCI-98"
/notes="end : 17"

```

ORIGIN

```

Query Match 76.0%; Score 19; DB 10; Length 1101;
Best Local Similarity 76.0%; Pred. No. 1.4e+03;
Matches 19; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

```

QY

```

1 AAAAAAATGACCAACTGGCGTGT 25
|||||:::|||||
450 AAAAAAAACCAACTGGCGTGT 474

```

RESULT 45

LOCUS BW779425/c

```

333 bp mRNA linear EST 10-AUG-2005
DEFINITION BW779425 Amphioxus Branchiostoma floridae unpublished cDNA library,
gastrula whole animal Branchiostoma floridae cDNA clone bbg0055a01
5', mRNA sequence.

```

ACCESSION

BW779425

VERSION BW779425.1

GI:66381919

KEYWORDS EST.

SOURCE Branchiostoma floridae (Florida lancelet)

ORGANISM Branchiostoma floridae

Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;

Branchiostoma.

1 (bases 1 to 333)

/notes="1 to 333"

AUTHORS Yu,J., Holland,L.Z., Shin-i,T., Kohara,Y., Satou,Y. and Satoh,N.

TITLE Expressed genes in Branchiostoma floridae

JOURNAL Unpublished (2005)

COMMENT Contact: Tadasu Shin-i

Center For Genetic Resource Information

National Institute of Genetics

1111 Yata, Mishima, Shizuoka 411-8540, Japan

Tel: 81-559-81-6856

Fax: 81-559-81-6855

Email: tshin@genes.nig.ac.jp

If you want to have a cDNA clone for this EST or if you have any

questions, please send an e-mail to Nori Satoh

(satoh@ascidian.zool.kyoto-u.ac.jp) and its cc to Yutaka Satou

(yutaka@ascidian.zool.kyoto-u.ac.jp).

FEATURES

source

```

Location/Qualifiers
1..333
/organism="Branchiostoma floridae"
/mol_type="mRNA"
/db_xref="taxon:7739"
/clone="bbg0055a01"
/tissue_type="whole animal"
/dev_stage="gastrula"
/clone_lib="Amphioxus Branchiostoma floridae unpublished
cDNA library, gastrula whole animal"

```

ORIGIN

```

Query Match 75.2%; Score 18.8; DB 5; Length 333;
Best Local Similarity 87.0%; Pred. No. 1.5e+03;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 1 AAAAAACATGACCAACTGGCT 23
Db 323 AAAAAACATGACCAACTGGCT 301

RESULT 46
AG0909802
LOCUS Drosophila auraria DNA, clone: DAB1-002H24.R.fa, genomic survey
sequence.
ACCESSION AG0909802
VERSION AG0909802.1 GI:58431114
KEYWORDS GSS
SOURCE Drosophila auraria
ORGANISM Drosophila auraria
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.

REFERENCE
1 Hattori,M., Toyoda,A., Murakami,K., Kuroki,Y., Fujiyama,A.,
Toshio,T.K. and Sakaki,Y.
BAC end sequences of Library DAB1
Unpublished
2 (bases 1 to 339)
Hattori,M.
Direct Submission
Submitted (19-JAN-2005) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa, 230-0045, Japan
(E-mail: hattori@gsc.riken.jp, URL:http://hgp.gsc.riken.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
Clones are derived from the BAC library DAB1
For BAC library availability, please contact Masa-Toshi Yamamoto
(yamamoto@kit.jp).
Submitted (30-11-2004) by Masahira Hattori,
RIKEN, Genomic Sciences Center (GSC);
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail: hattori@gsc.riken.jp, Tel:81-45-503-9111,
Fax:81-45-503-9170)
This work was done in collaboration with Yamamoto, M.-T. Drosophila
Genetic Resource Center
Saga Ippongi-cho, Ukyo-ku, Kyoto 616-8354, Japan
Tel: 81-75-873-2660 FAX: 81-75-861-0881
PRIMERS
Sequencing : R

LIBRARY
Vector : pKS150
R.Site 1 : SacI
R.Site 2 : SacI.
Location/Qualifiers
1. .339
/organism="Drosophila auraria"
/mol_type="genomic DNA"
/db_xref="taxon:47315"
/clone="DAB1-002H24.R.fa"
/clone_lib="DAB1 Drosophila BAC library"

ORIGIN
Query Match 75.2%; Score 18.8; DB 10; Length 339;
Best Local Similarity 90.9%; Pred. No. 1.5e+03;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AAAAAACATGACCAACTGGCT 23
Db 152 AAAAAACATGACCAACTGGCT 173

RESULT 47
BY066857
LOCUS RIKEN full-length enriched, 17 days embryo stomach Mus
DEFINITION musculus cDNA clone 1920051C22 5', mRNA sequence.
ACCESSION BY066857

VERSION
KEYWORDS
SOURCE
ORGANISM
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Murinae; Mus.

1 (bases 1 to 410)
Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S.,
Nikaido,I., Osato,N., Saito,R., Suzuki,H., Yamada,K., I.,
Kiyosawa,H., Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A.,
Schonbach,C., Gojobori,T., Baldarelli,R., Hill,D.P., Bult,C.,
Hume,D.A., Quackenbush,J., Schriml,L.M., Kanapin,A., Matsuda,H.,
Batalov,S., Beisel,K.W., Blake,J.A., Bradt,D., Brusic,V.,
Chothia,C., Corbani,L.E., Cousins,S., Della,E., Dragani,T.A.,
Fletcher,C.F., Forrest,A., Frazer,K.S., Gaasterland,T.,
Gariboldi,M., Gissi,C., Godzik,A., Gough,J., Grimmond,S.,
Gustincich,S., Hirokawa,N., Jackson,I.J., Jarvis,E.D., Kanai,A.,
Kawaji,H., Kawasawa,Y., Kedzierski,R.M., King,B.L., Konegaya,A.,
Kurochin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R.,
Maltais,L., Marchionni,L., McKensie,L., Miki,H., Nagashima,T.,
Numata,K., Okido,I., Pavan,W.J., Pertea,G., Pesole,G.,
Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramachandran,S.,
Ravasi,T., Reed,J.C., Reid,D.J., Reid,J., Ring,B.Z., Ringwald,M.,
Sandelin,A., Schneider,C., Sempile,C.A., Setou,M., Shimada,K.,
Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tomita,M.,
Verardo,C., Wagner,L., Wahlestedt,C., Wang,Y., Watanabe,Y.,
Wells,C., Wilming,L.G., Wynshaw-Boris,A., Yanagisawa,M., Yang,I.,
Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A., Carninci,P.,
Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura,M.,
Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,K., Aizawa,K.,
Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K., Ishii,Y.,
Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata,K.,
Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,K., Lander,E.S.,
Rogers,J., Birney,E. and Hayashizaki,Y.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
12466851
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/
Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S.,
Hirozane,T., Imotani,K., Ishii,Y., Itoh,M., Kawai,J., Konno,H.,
Miyazaki,A., Murata,M., Nakamura,M., Nomura,K., Numazaki,R.,
Ohno,M., Sakai,K., Sakazume,N., Sasaki,D., Sato,K., Shibata,K.,
Shiraki,T., Tagami,M., Waki,K., Watahiki,A., Muramatsu,M. and
Hayashizaki,Y. Direct Submission
Computational Analysis of Full-Length Mouse cDNAs Compared with
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site (http://genome.gsc.riken.go.jp) for
further details.
Location/Qualifiers
1. .410

FEATURES
source

```

```

/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="1920051C22"
/tissue_type="stomach"
/dev_stage="17 days embryo"
/clone_lib="RIKEN full-length enriched, 17 days embryo
stomach"

ORIGIN
Query Match          75.2%; Score 18.8; DB 5; Length 410;
Best Local Similarity 90.9%; Pred. No. 1.5e+03;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 AACAAATGACCAACTGGGCTGT 25
||||| ||||| ||||| ||||| |||||
Db 357 AACAAAGGACCAATGGGCTGT 378

RESULT 48
AQ736236          511 bp      DNA      linear      GSS 15-JUL-1999
LOCUS
DEFINITION
HS 3043 B2 P05 MR CIT Approved Human Genomic Sperm Library D Homo
sapiens genomic clone Plate=3043 Col=10 Row=H, genomic survey
sequence.
ACCESSION
AQ736236
VERSION
AQ736236.1 GI:5507788
KEYWORDS
GSS.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE
1 (bases 1 to 511)
Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
Hood,L.
Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
1049764
Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones may be purchased from Research Genetics (info@resgen.com).
BAC end Web Server: http://www.htsc.washington.edu
Plate: 3043 row: H column: 10
Seq primer: M13 Reverse
Class: BAC ends
High quality sequence stop: 511.
FEATURES
source
1..511
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="Plate=3043 Col=10 Row=H"
/sex="male"
/clone_lib="CIT Approved Human Genomic Sperm Library D"
/note="Organ: sperm; Vector: pBelBAC11; BAC Clones in
E-Coli DH10B"

ORIGIN
Query Match          75.2%; Score 18.8; DB 9; Length 511;
Best Local Similarity 90.9%; Pred. No. 1.6e+03;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAACAATGACCAACTGGGC 22
||||| ||||| ||||| ||||| |||||

```

```

Db 150 AAAAATAATGACCACTGGGC 171

RESULT 49
CB334117/c
LOCUS
DEFINITION
CB334117          523 bp      mRNA      linear      EST 29-OCT-2003
Mg_AFT_30D09_M13F Mesobuthus gibbosus adult female tail Mesobuthus
gibbosus cDNA clone Mg_AFT_30D09 5' similar to P91468 (P91468)
T20D4.7 protein, mRNA sequence.
ACCESSION
CB334117
VERSION
CB334117.1 GI:28915869
KEYWORDS
EST.
SOURCE
Mesobuthus gibbosus (Mediterranean checkered scorpion)
ORGANISM
Mesobuthus gibbosus
Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
Buthida; Buthoidea; Buthidae; Mesobuthus.
REFERENCE
1 (bases 1 to 523)
Gantenbein,B., Thomson,M., Rosie,A., Parkinson,J. and Blaxter,M.
Gene discovery and phylogenetics of chelicerate arthropods using
expressed sequence tags
Unpublished (2002)
JOURNAL
Contact: Gantenbein-Ritter B
Institute of Cell, Animal and Population Biology
University of Edinburgh
West Mains Rd, Edinburgh, UK
Email: B.Gantenbein@ed.ac.uk
The library was prepared by Benjamin Gantenbein-Ritter and Marian
Thomson, University of Edinburgh, from the Metasoma + Telson of an
adult female. Sequencing was performed in Mark Blaxters lab in
Edinburgh.
PCR Primers
FORWARD: M13F
BACKWARD: pDNRlib_M13R
Plate: 30 row: D column: 09
Seq primer: M13F
High quality sequence stop: 370.
FEATURES
source
1..523
Location/Qualifiers
/organism="Mesobuthus gibbosus"
/mol_type="mRNA"
/strain="Farga"
/db_xref="taxon:123226"
/clone="Mg_AFT_30D09"
/sex="female"
/tissue_type="Metasoma + Telson"
/dev_stage="adult"
/clone_lib="Mesobuthus gibbosus adult female tail"
/note="The Mesobuthus gibbosus EST library was prepared by
Benjamin Gantenbein and Marian Thomson, University of
Edinburgh. cDNA from the metasoma and telson of an adult
female was cloned into the vector pDNR-LIB. ESTs were
amplified from each clone by PCR, using primers pDNR.M13F
and pDNR.M13R. The products were cleaned of excess
nucleotides and phosphates, by treatment with SAP/ExoI,
and then sequenced using pDNR.seq primer."
ORIGIN
Query Match          75.2%; Score 18.8; DB 6; Length 523;
Best Local Similarity 90.9%; Pred. No. 1.6e+03;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 AAAACAATGACCAACTGGGCTG 24
||||| ||||| ||||| ||||| |||||
Db 43 AAAAATAATGTCCAACTGGGCTG 22

RESULT 50
AG914898
LOCUS
DEFINITION
AG914898          687 bp      DNA      linear      GSS 01-FEB-2005
Drosophila auraria DNA, clone: DAB1-011L05.R.fa, genomic survey
sequence.
ACCESSION
AG914898
VERSION
AG914898.1 GI:58436254

```

KEYWORDS GSS.
SOURCE Drosophila auraria
ORGANISM Drosophila auraria
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1
AUTHORS Hattori, M., Toyoda, A., Murakami, K., Kuroki, Y., Fujiyama, A.,
Toshio, T. K. and Sakaki, Y.
TITLE BAC end sequences of Library DAB1
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 687)
AUTHORS Hattori, M.
TITLE Direct Submission
JOURNAL Submitted (19-JAN-2005) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa, 230-0045, Japan
(E-mail: hattori@gsc.riken.jp. URL: http://hgp.gsc.riken.jp/,
Tel: 81-45-503-9111, Fax: 81-45-503-9170)
COMMENT Clones are derived from the BAC library DAB1
For BAC library availability, please contact Masa-Toshi Yamamoto
(yamamoto@kit.jp).
Submitted (30-11-2004) by Masahira Hattori,
RIKEN, Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail: hattori@gsc.riken.jp, tel: 81-45-503-9111,
Fax: 81-45-503-9170)
This work was done in collaboration with Yamamoto, M.-T. Drosophila
Genetic Resource Center
Saga Ippongi-cho, Ukyo-ku, Kyoto 616-8354, Japan
Tel: 81-75-873-2660 FAX: 81-75-861-0881
PRIMERS
Sequencing : R
LIBRARY
Vector : pKSI50
R.Site 1 : SacI
R.Site 2 : SacI
Location/Qualifiers
1. .687
/organism="Drosophila auraria"
/mol_type="genomic DNA"
/db_xref="taxon:47315"
/clone="DAB1-011L05.R.fa"
/clone_lib="DAB1 Drosophila BAC library"
ORIGIN
Query Match 75.2%; Score 18.8; DB 10; Length 687;
Best Local Similarity 90.9%; Pred. No. 1.6e+03;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 2 AAAAAACAATGACCAACTGGGCT 23
|||||
Db 273 AAAAAACATTGACCAACTGGTCT 294
Search completed: February 3, 2006, 22:01:55
Job time : 2961.67 secs

GenCore version 5.1.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 3, 2006, 11:35:31 ; Search time 80.6667 Seconds
(without alignments)
550.897 Million cell updates/sec

Title: US-10-719-900-5

Perfect score: 25

Sequence: 1 aaaaacaatgacgaactgggctgt 25

Scoring table: IDENTIFY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database : Issued Patents NA:*

- 1: /cgn2_6/ptodata/1/ina/1 COMB.seq:*
- 2: /cgn2_6/ptodata/1/ina/5 COMB.seq:*
- 3: /cgn2_6/ptodata/1/ina/6A COMB.seq:*
- 4: /cgn2_6/ptodata/1/ina/6B COMB.seq:*
- 5: /cgn2_6/ptodata/1/ina/H COMB.seq:*
- 6: /cgn2_6/ptodata/1/ina/PCTUS COMB.seq:*
- 7: /cgn2_6/ptodata/1/ina/PP COMB.seq:*
- 8: /cgn2_6/ptodata/1/ina/RE COMB.seq:*
- 9: /cgn2_6/ptodata/1/ina/backfileseq.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	22.4	89.6	601	3	US-09-949-016-128244
2	20.8	83.2	601	3	US-09-949-016-128243
C 3	20.8	83.2	11994	3	US-09-949-016-15373
C 4	18.8	75.2	26843	3	US-09-949-016-17208
5	18.6	74.4	87523	3	US-09-949-016-12670
6	18.6	74.4	87523	3	US-09-949-016-15047
7	18.6	74.4	87523	3	US-09-949-016-15048
8	18.6	74.4	87523	3	US-09-949-016-15049
9	18.6	74.4	87869	3	US-09-949-016-11744
10	18.6	74.4	87869	3	US-09-949-016-15044
11	18.6	74.4	87869	3	US-09-949-016-15045
12	18.6	74.4	87869	3	US-09-949-016-15046
C 13	18.4	73.6	601	3	US-09-949-016-191078
C 14	17.6	70.4	601	3	US-09-949-016-40234
C 15	17.6	70.4	601	3	US-09-949-016-51844
C 16	17.6	70.4	601	3	US-09-949-016-139379
C 17	17.6	70.4	601	3	US-09-949-016-154598
C 18	17.6	70.4	3256	3	US-10-104-047-474
C 19	17.6	70.4	13335	3	US-09-949-016-13162
C 20	17.6	70.4	80269	3	US-09-949-016-15681
21	17.6	70.4	132456	3	US-09-949-016-13750
22	17.6	70.4	146095	3	US-09-949-016-12872
23	17.6	70.4	146104	3	US-09-949-016-13239
24	17.6	70.4	275110	3	US-09-949-016-12706

25	17.6	70.4	275110	3	US-09-949-016-16070	Sequence 16070, A
C 26	17.2	68.8	601	3	US-09-949-016-129135	Sequence 129135, A
27	17.2	68.8	3381	3	US-09-009-119-1	Sequence 1, Appli
28	17.2	68.8	3381	3	US-09-371-507-1	Sequence 1, Appli
29	17.2	68.8	3383	6	PCT-US95-09098-1	Sequence 1, Appli
30	17.2	68.8	14079	3	US-09-949-016-11993	Sequence 11993, A
C 31	17.2	68.8	98864	3	US-09-949-016-15403	Sequence 15403, A
C 32	17	68.0	500	3	US-10-002-623-908	Sequence 908, App
C 33	17	68.0	601	3	US-09-949-016-133748	Sequence 133748, A
C 34	17	68.0	601	3	US-09-949-016-133772	Sequence 133772, A
35	17	68.0	601	3	US-09-949-016-151005	Sequence 151005, A
C 36	17	68.0	601	3	US-09-949-016-169748	Sequence 169748, A
C 37	17	68.0	601	3	US-09-949-016-137059	Sequence 137059, A
C 38	17	68.0	700	3	US-09-735-271-884	Sequence 884, App
C 39	17	68.0	8888	3	US-09-949-016-12308	Sequence 12308, A
C 40	17	68.0	8889	3	US-09-949-016-15760	Sequence 15760, A
41	17	68.0	15585	3	US-09-949-016-11927	Sequence 11927, A
42	17	68.0	15585	3	US-09-949-016-15627	Sequence 15627, A
43	17	68.0	36023	3	US-09-949-016-15577	Sequence 15577, A
C 44	17	68.0	45842	3	US-09-949-016-12550	Sequence 12550, A
C 45	17	68.0	45842	3	US-09-949-016-17327	Sequence 17327, A
46	17	68.0	56241	3	US-09-949-016-15174	Sequence 15174, A
47	17	68.0	56241	3	US-09-949-016-15175	Sequence 15175, A
C 48	17	68.0	75295	3	US-09-949-002-575	Sequence 575, App
C 49	17	68.0	75296	3	US-09-949-002-799	Sequence 799, App
50	17	68.0	99748	3	US-09-949-016-11990	Sequence 11990, A
51	17	68.0	99749	3	US-09-949-016-16518	Sequence 16518, A
52	17	68.0	116652	3	US-09-949-016-13413	Sequence 13413, A
53	17	68.0	183770	3	US-09-949-016-15494	Sequence 15494, A
54	17	68.0	239527	3	US-09-949-016-15980	Sequence 15980, A
C 55	17	68.0	323820	3	US-09-949-016-14139	Sequence 14139, A
C 56	17	68.0	421491	3	US-09-949-016-12805	Sequence 12805, A
57	17	68.0	421494	3	US-09-949-016-14060	Sequence 14060, A
58	16.8	67.2	652	3	US-09-270-767-4915	Sequence 4915, Ap
59	16.8	67.2	652	3	US-09-270-767-20197	Sequence 20197, A
60	16.8	67.2	1764	3	US-09-579-383-2	Sequence 2, Appli
C 61	16.8	67.2	2500	3	US-09-579-383-5	Sequence 5, Appli
C 62	16.8	67.2	193555	3	US-09-949-016-15553	Sequence 15553, A
C 63	16.8	67.2	193555	3	US-09-949-016-15554	Sequence 15554, A
C 64	16.8	67.2	193555	3	US-09-949-016-15555	Sequence 15555, A
C 65	16.6	66.4	396	3	US-09-583-110-234	Sequence 234, App
C 66	16.6	66.4	601	3	US-09-949-016-68220	Sequence 68220, A
C 67	16.6	66.4	601	3	US-09-949-016-68221	Sequence 68221, A
C 68	16.6	66.4	601	3	US-09-949-016-68222	Sequence 68222, A
C 69	16.6	66.4	601	3	US-09-949-016-76457	Sequence 76457, A
70	16.6	66.4	601	3	US-09-949-016-85312	Sequence 85312, A
71	16.6	66.4	1917	3	US-09-270-767-15042	Sequence 15042, A
72	16.6	66.4	9279	3	US-09-487-558B-171	Sequence 171, App
C 73	16.6	66.4	10764	3	US-09-949-016-11872	Sequence 11872, A
C 74	16.6	66.4	10765	3	US-09-949-016-14204	Sequence 14204, A
C 75	16.6	66.4	11378	3	US-08-961-527-210	Sequence 210, App
C 76	16.6	66.4	31861	3	US-09-949-016-12803	Sequence 12803, A
C 77	16.6	66.4	31861	3	US-09-949-016-13967	Sequence 13967, A
C 78	16.6	66.4	37288	3	US-09-949-016-14593	Sequence 14593, A
C 79	16.6	66.4	37288	3	US-09-949-016-14594	Sequence 14594, A
C 80	16.6	66.4	71863	3	US-09-949-016-15112	Sequence 15112, A
C 81	16.6	66.4	83851	3	US-09-949-016-13847	Sequence 13847, A
C 82	16.6	66.4	88557	3	US-09-949-016-17028	Sequence 17028, A
C 83	16.6	66.4	94830	3	US-09-949-016-12414	Sequence 12414, A
C 84	16.6	66.4	94847	3	US-09-949-016-16336	Sequence 16336, A
C 85	16.6	66.4	115954	3	US-09-949-016-12298	Sequence 12298, A
C 86	16.6	66.4	119594	3	US-09-949-016-12080	Sequence 12080, A
C 87	16.6	66.4	119594	3	US-09-949-016-15952	Sequence 15952, A
C 88	16.6	66.4	119601	3	US-09-949-016-14711	Sequence 14711, A
C 89	16.6	66.4	156324	3	US-09-949-016-13749	Sequence 13749, A
C 90	16.6	66.4	194933	3	US-09-949-016-14172	Sequence 14172, A
C 91	16.6	66.4	200663	3	US-09-949-016-12569	Sequence 12569, A
C 92	16.6	66.4	213456	3	US-09-820-007-3	Sequence 3, Appli
C 93	16.6	66.4	294816	3	US-09-949-016-15974	Sequence 15974, A
94	16.6	66.4	330416	3	US-09-949-016-16923	Sequence 16923, A
95	16.4	65.6	1490	3	US-10-200-012-34	Sequence 34, Appli
96	16.4	65.6	1652	3	US-09-774-528-17	Sequence 17, Appli
97	16.4	65.6	1652	3	US-10-120-988-17	Sequence 17, Appli

```

98      16.4      65.6      1818      3      US-10-104-047-670      Sequence 670, App
c 99      16.2      64.8      306      2      US-08-634-797-17      Sequence 17, Appl
c 100     16.2      64.8      333      2      US-08-634-797-21      Sequence 21, Appl
c 101     16.2      64.8      333      2      US-08-634-797-31      Sequence 31, Appl
c 102     16.2      64.8      378      3      US-09-513-999C-19527      Sequence 19527, A
c 103     16.2      64.8      576      2      US-08-086-428B-26      Sequence 26, Appl
c 104     16.2      64.8      576      2      US-08-468-570-26      Sequence 26, Appl
c 105     16.2      64.8      576      2      US-08-290-665A-26      Sequence 26, Appl
c 106     16.2      64.8      576      3      US-08-466-601A-26      Sequence 26, Appl
c 107     16.2      64.8      576      6      PCT-US95-10398-26      Sequence 26, Appl
c 108     16.2      64.8      601      3      US-09-949-016-79047      Sequence 79047, A
c 109     16.2      64.8      601      3      US-09-949-016-146997      Sequence 146997,
c 110     16.2      64.8      601      3      US-09-949-016-152677      Sequence 152677,
c 111     16.2      64.8      601      3      US-09-949-016-152678      Sequence 152678,
c 112     16.2      64.8      1616      3      US-09-907-794A-113      Sequence 113, App
c 113     16.2      64.8      1616      3      US-09-905-128A-113      Sequence 113, App
c 114     16.2      64.8      1616      3      US-09-902-775A-113      Sequence 113, App
c 115     16.2      64.8      1616      3      US-09-906-700-113      Sequence 113, App
c 116     16.2      64.8      1616      3      US-09-903-603A-113      Sequence 113, App
c 117     16.2      64.8      1616      3      US-09-904-920A-113      Sequence 113, App
c 118     16.2      64.8      1616      3      US-09-909-064-113      Sequence 113, App
c 119     16.2      64.8      1616      3      US-09-905-381A-113      Sequence 113, App
c 120     16.2      64.8      1616      3      US-09-906-618-113      Sequence 113, App
c 121     16.2      64.8      1616      3      US-09-906-646-113      Sequence 113, App
c 122     16.2      64.8      1616      3      US-09-904-463-113      Sequence 113, App
c 123     16.2      64.8      1616      3      US-09-902-736A-113      Sequence 113, App
c 124     16.2      64.8      1616      3      US-09-906-722A-113      Sequence 113, App
c 125     16.2      64.8      6921      3      US-09-643-597-117      Sequence 117, App
c 126     16.2      64.8      6921      3      US-09-480-884A-117      Sequence 117, App
c 127     16.2      64.8      6921      3      US-09-542-615A-117      Sequence 117, App
c 128     16.2      64.8      6921      3      US-09-606-421B-117      Sequence 117, App
c 129     16.2      64.8      6921      3      US-09-221-107-117      Sequence 117, App
c 130     16.2      64.8      6921      3      US-09-466-396A-117      Sequence 117, App
c 131     16.2      64.8      6921      3      US-09-476-496A-117      Sequence 117, App
c 132     16.2      64.8      6921      3      US-09-630-940B-117      Sequence 117, App
c 133     16.2      64.8      6921      3      US-09-285-479-117      Sequence 117, App
c 134     16.2      64.8      6921      3      US-10-007-700-117      Sequence 117, App
c 135     16.2      64.8      7045      3      US-09-919-172-28      Sequence 28, Appl
c 136     16.2      64.8      8934      3      US-09-949-016-11903      Sequence 11903, A
c 137     16.2      64.8      8930      3      US-09-814-915A-91      Sequence 91, Appl
c 138     16.2      64.8      9589      2      US-07-925-695-1        Sequence 1, Appl
c 139     16.2      64.8      9589      2      US-07-925-695-2        Sequence 2, Appl
c 140     16.2      64.8      36820      3      US-09-949-016-16665      Sequence 16665, A
c 141     16.2      64.8      46085      3      US-09-949-016-13547      Sequence 13547, A
c 142     16.2      64.8      46085      3      US-09-949-016-13548      Sequence 13548, A
c 143     16.2      64.8      81461      3      US-09-949-016-16419      Sequence 16419, A
c 144     16.2      64.8      83708      3      US-09-949-016-17207      Sequence 17207, A
c 145     16.2      64.8      87562      3      US-09-949-016-13685      Sequence 13685, A
c 146     16.2      64.8      114793      3      US-10-148-806-3        Sequence 3, Appl
c 147     16.2      64.8      155019      3      US-09-949-016-16029      Sequence 16029, A
c 148     16.2      64.8      374159      3      US-09-949-016-15868      Sequence 15868, A
c 149     16.2      64.8      784019      3      US-09-949-016-14033      Sequence 14033, A
c 150     16.2      64.8      828152      3      US-09-949-016-12777      Sequence 12777, A

ALIGNMENTS

RESULT 1
US-09-949-016-128244
; Sequence 128244, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 128244
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-128243
; Sequence 128243, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 128243
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-128243
Query Match      83.2%; Score 20.8; DB 3; Length 601;
Best Local Similarity 91.7%; Pred. NO. 6.6;
Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1      AAAAAACAATGACCAACTGGGCTG 24
Db      21      AAAAAACAATGTCCAACTGGTCTG 44

RESULT 3
US-09-949-016-15373/c
; Sequence 15373, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15373
```

; LENGTH: 11994
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15373

Query Match 83.2%; Score 20.8; DB 3; Length 11994;
Best Local Similarity 91.7%; Pred. No. 9.7;
Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAACAATGACCAACTGGGCTG 24
DB 8548 AAAAAACAATGTCACCACTGGGCTG 8525

RESULT 4

US-09-949-016-17208/c
; Sequence 17208, Application US/09949016
; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17208
; LENGTH: 26843
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17208

Query Match 75.2%; Score 18.8; DB 3; Length 26843;
Best Local Similarity 90.9%; Pred. No. 85;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AAAAAACAATGACCAACTGGGCT 23
DB 13095 AAAAAACAATGACCACTGGGCT 13074

RESULT 5

US-09-949-016-12670
; Sequence 12670, Application US/09949016
; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12670
; LENGTH: 87523
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12670

Query Match 74.4%; Score 18.6; DB 3; Length 87523;
Best Local Similarity 84.0%; Pred. No. 1.2e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAACAATGACCAACTGGGCTGT 25
DB 6500 AAAAAACAATCACCACCAATGGGATGT 6524

RESULT 6

US-09-949-016-15047
; Sequence 15047, Application US/09949016
; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15047
; LENGTH: 87523
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15047

Query Match 74.4%; Score 18.6; DB 3; Length 87523;
Best Local Similarity 84.0%; Pred. No. 1.2e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAACAATGACCAACTGGGCTGT 25
DB 6500 AAAAAACAATCACCACCAATGGGATGT 6524

RESULT 7

US-09-949-016-15048
; Sequence 15048, Application US/09949016
; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15048
; LENGTH: 87523
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15048

Query Match 74.4%; Score 18.6; DB 3; Length 87523;
Best Local Similarity 84.0%; Pred. No. 1.2e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAACAATGACCAACTGGGCTGT 25

```
Db      6500 AAAAAAATACCAAAATGGGATGT 6524
||||| ||| ||| ||| ||| ||| ||| |||
US-09-949-016-15049
; Sequence 15049, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15049
; LENGTH: 87523
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15049

Query Match      74.4%; Score 18.6; DB 3; Length 87523;
Best Local Similarity 84.0%; Pred. No. 1.2e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      1 AAAAAACAATGACCAACTGGGCTGT 25
||||| ||| ||| ||| ||| ||| |||
Db      6500 AAAAAAATACCAAAATGGGATGT 6524

RESULT 9
US-09-949-016-11744
; Sequence 11744, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11744
; LENGTH: 87869
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-11744

Query Match      74.4%; Score 18.6; DB 3; Length 87869;
Best Local Similarity 84.0%; Pred. No. 1.2e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      1 AAAAAACAATGACCAACTGGGCTGT 25
||||| ||| ||| ||| ||| ||| |||
Db      6846 AAAAAAATACCAAAATGGGATGT 6870

RESULT 10
US-09-949-016-15046
; Sequence 15046, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
```

```
US-09-949-016-15044
; Sequence 15044, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15044
; LENGTH: 87869
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15044

Query Match      74.4%; Score 18.6; DB 3; Length 87869;
Best Local Similarity 84.0%; Pred. No. 1.2e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      1 AAAAAACAATGACCAACTGGGCTGT 25
||||| ||| ||| ||| ||| ||| |||
Db      6846 AAAAAAATACCAAAATGGGATGT 6870

RESULT 11
US-09-949-016-15045
; Sequence 15045, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15045
; LENGTH: 87869
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15045

Query Match      74.4%; Score 18.6; DB 3; Length 87869;
Best Local Similarity 84.0%; Pred. No. 1.2e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      1 AAAAAACAATGACCAACTGGGCTGT 25
||||| ||| ||| ||| ||| ||| |||
Db      6846 AAAAAAATACCAAAATGGGATGT 6870

RESULT 12
US-09-949-016-15046
; Sequence 15046, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
```

```
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15046
; LENGTH: 87869
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15046

Query Match          74.4%; Score 18.6; DB 3; Length 87869;
Best Local Similarity 84.0%; Pred. No. 1.2e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAACAATGACCAACTGGGCTGT 25
Db 6846 AAAAAAATACCAAAATGGGATGT 6870

RESULT 13
US-09-949-016-191078/c
; Sequence 191078, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 191078
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-191078

Query Match          73.6%; Score 18.4; DB 3; Length 601;
Best Local Similarity 86.4%; Pred. No. 78;
Matches 19; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 AAAAAACAATGACCAACTGGGCT 23
Db 306 AAAAAAAGACCACTGGGCT 285

RESULT 14
US-09-949-016-40234/c
; Sequence 40234, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
```

```
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 40234
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-40234

Query Match          70.4%; Score 17.6; DB 3; Length 601;
Best Local Similarity 83.3%; Pred. No. 1.8e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAACAATGACCAACTGGGCTG 24
Db 253 AAAAAACAATGACCAACCGCCTG 230

RESULT 15
US-09-949-016-51844/c
; Sequence 51844, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 51844
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-51844

Query Match          70.4%; Score 17.6; DB 3; Length 601;
Best Local Similarity 83.3%; Pred. No. 1.8e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAACAATGACCAACTGGGCTG 24
Db 253 AAAAAACAATGACCAACCGCCTG 230

RESULT 16
US-09-949-016-139379
; Sequence 139379, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
```

```
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 139379
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-139379

Query Match          70.4%; Score 17.6; DB 3; Length 601;
Best Local Similarity 83.3%; Pred. No. 1.8e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AAAACAATGACCAACTGGGCTGT 25
    ||||| ||||| ||||| |||||
Db 560 AAAACAATGAACAGCTAAGCTGT 583

RESULT 17
US-09-949-016-154598/c
; Sequence 154598, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 154598
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-154598

Query Match          70.4%; Score 17.6; DB 3; Length 601;
Best Local Similarity 83.3%; Pred. No. 1.8e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AAAACAATGACCAACTGGGCTGT 25
    ||||| ||||| ||||| |||||
Db 117 AAAACAAGACAACTGGATTGT 94

RESULT 18
US-10-104-047-474
; Sequence 474, Application US/10104047
; Patent No. 6943241
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. 6943241el full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 474
; LENGTH: 3256
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-104-047-474

Query Match          70.4%; Score 17.6; DB 3; Length 3256;
Best Local Similarity 83.3%; Pred. No. 1.8e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AAAACAATGACCAACTGGGCTGT 25
    ||||| ||||| ||||| |||||
Db 117 AAAACAAGACAACTGGATTGT 94

RESULT 19
US-09-949-016-13162/c
; Sequence 13162, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13162
; LENGTH: 13335
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13162

Query Match          70.4%; Score 17.6; DB 3; Length 13335;
Best Local Similarity 83.3%; Pred. No. 2.7e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAACAATGACCAACTGGGCTG 24
    ||||| ||||| ||||| |||||
Db 3954 AAAAAAAGCAACTGGGCTG 3931

RESULT 20
US-09-949-016-15681
; Sequence 15681, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15681
; LENGTH: 80269
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15681

Query Match          70.4%; Score 17.6; DB 3; Length 80269;
Best Local Similarity 83.3%; Pred. No. 3.4e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AAAACAATGACCAACTGGGCTGT 25
    ||||| ||||| ||||| |||||
```

Db 73866 AAAAAACAATGACACAGCTAAGCTGT 73889

RESULT 21

US-09-949-016-13750
; Sequence 13750, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13750
; LENGTH: 132456
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(132456)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13750

Query Match 70.4%; Score 17.6; DB 3; Length 132456;
Best Local Similarity 83.3%; Pred. No. 3.6e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AAAAAACAATGACCAACTGGGCTGT 25

Db 47727 AAACACAATGACCAAGTGACTTT 47750

RESULT 22

US-09-949-016-12872
; Sequence 12872, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12872
; LENGTH: 146095
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12872

Query Match 70.4%; Score 17.6; DB 3; Length 146095;
Best Local Similarity 83.3%; Pred. No. 3.6e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAACAATGACCAACTGGGCTG 24

Db 140538 AAAAAACAATGACCAACGCGCTG 140561

RESULT 23

US-09-949-016-13239
; Sequence 13239, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13239
; LENGTH: 146104
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13239

Query Match 70.4%; Score 17.6; DB 3; Length 146104;
Best Local Similarity 83.3%; Pred. No. 3.6e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAACAATGACCAACTGGGCTG 24

Db 140538 AAAAAACAATGACCAACGCGCTG 140561

RESULT 24

US-09-949-016-12706
; Sequence 12706, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12706
; LENGTH: 275110
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12706

Query Match 70.4%; Score 17.6; DB 3; Length 275110;
Best Local Similarity 83.3%; Pred. No. 3.9e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AAAAAACAATGACCAACTGGGCTGT 25

Db 27732 AAAAAACAAGACAAACTGGATTGT 27755

RESULT 25

US-09-949-016-16070
; Sequence 16070, Application US/09949016

```

; Patent No. 6812339
;
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CU001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: fastSEQ for Windows Version 4.0
; SEQ ID NO 16070
; LENGTH: 275110
; TYPE: DNA
; ORGANISM: Human
; PS-09-949-016-16070

```

```

; APPLICANT: HARRIS, Elizabeth H.
; TITLE OF INVENTION: Porphyrin Accumulating-Type Herbicide Resistance Gene
; FILE REFERENCE: substitute sequence listing
; CURRENT APPLICATION NUMBER: US/09/009,119
; CURRENT FILING DATE: 1998-01-20
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 3381
; TYPE: DNA
; ORGANISM: Chlamydomonas reinhardtii
US-09-009-119-1

Query Match          68.8%; Score 17.2; DB 3; Length 3381;
Best Local Similarity 86.4%; Pred. No. 3.4e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy      4 AAACAATGCACCAACTGGGCTGT 25
        ||||||| ||| |||||
Db      340 AAACAATGCACAAACGGGGCTGT 361

RESULT 28
US-09-371-507-1
; Sequence 1, Application US/09371507
; Patent No. 6346656
; GENERAL INFORMATION:
; APPLICANT: SATO, Ryo
; APPLICANT: BOYNTON, John E.
; APPLICANT: GILLHAM, Nicholas W.
; APPLICANT: HARRIS, Elizabeth H.
; TITLE OF INVENTION: Porphyrin Accumulating-Type Herbicide Resistance Gene
; FILE REFERENCE: substitute sequence listing
; CURRENT APPLICATION NUMBER: US/09/371,507
; CURRENT FILING DATE: 1999-08-18
; PRIOR APPLICATION NUMBER: 09/009,119
; PRIOR FILING DATE: 1998-01-20
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 3381
; TYPE: DNA
; ORGANISM: Chlamydomonas reinhardtii
US-09-371-507-1

Query Match          68.8%; Score 17.2; DB 3; Length 3381;
Best Local Similarity 86.4%; Pred. No. 3.4e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy      4 AAACAATGCACCAACTGGGCTGT 25
        ||||||| ||| |||||
Db      340 AAACAATGCACACGGGGCTGT 361

RESULT 29
PCT-US95-09098-1
; Sequence 1, Application PC/TUS9509098
; GENERAL INFORMATION:
; APPLICANT: Sato, Ryo
; APPLICANT: Boynton, John E.
; APPLICANT: Gillham, Nicholas W.
; APPLICANT: Harris, Elizabeth H.
; TITLE OF INVENTION: Porphyrin-Accumulating Type Herbicide
; TITLE OF INVENTION: Resistance Gene
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch, LLP
; STREET: 8110 Gatehouse Road, Suite 500 East
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22042
; COMPUTER READABLE FORM:

```


MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/09098
FILING DATE: 20-JUL-1995
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Murphy Jr., Gerald M.
REGISTRATION NUMBER: 28,977
REFERENCE/DOCKET NUMBER: 2185-110P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
TELEX: 248345
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3383 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Chlamydomonas reinhardtii
STRAIN: RS-3
PCT-US95-09098-1

Query Match 68.8%; Score 17.2; DB 6; Length 3383;
Best Local Similarity 86.4%; Pred. No. 3.4e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 AACAATGACCAACTGGGCTGT 25
||||| ||||| ||||| |||||
Db 340 AACAATGCACACAGGGGCTGT 361

RESULT 30
US-09-949-016-11993
; Sequence 11993, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11993
; LENGTH: 14079
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-11993

Query Match 68.8%; Score 17.2; DB 3; Length 14079;
Best Local Similarity 86.4%; Pred. No. 4.1e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 AACAATGACCAACTGGGCTG 24
||||| ||||| ||||| |||||
Db 358 AACAATGGCAACTGGGCTG 379

RESULT 31
US-09-949-016-15403/c
; Sequence 15403, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15403
; LENGTH: 98864
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(98864)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15403

Query Match 68.8%; Score 17.2; DB 3; Length 98864;
Best Local Similarity 86.4%; Pred. No. 5.2e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAAATGACCAACTGGGC 22
||||| ||||| ||||| |||||
Db 24111 AATAAAAAAGACCAACTGGGC 24090

RESULT 32
US-10-002-623-908/c
; Sequence 908, Application US/10002623
; Patent No. 6929911
; GENERAL INFORMATION:
; APPLICANT: OEFNER, PETER J.
; APPLICANT: UNDERHILL, PETER A.
; TITLE OF INVENTION: A METHOD FOR DETERMINING GENETIC
; TITLE OF INVENTION: AFFILIATION, SUBSTRUCTURE AND GENE FLOW WITHIN HUMAN
; TITLE OF INVENTION: POPULATIONS
; FILE REFERENCE: STAN-212
; CURRENT APPLICATION NUMBER: US/10/002,623
; CURRENT FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: US 60/245,355
; PRIOR FILING DATE: 2000-11-01
; NUMBER OF SEQ ID NOS: 952
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 908
; LENGTH: 500
; TYPE: DNA
; ORGANISM: Homo Sapiens
; US-10-002-623-908

Query Match 68.0%; Score 17; DB 3; Length 500;
Best Local Similarity 80.0%; Pred. No. 3.2e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAAAATGACCAACTGGGCTG 25
||||| ||||| ||||| |||||
Db 397 AAAAAAAGAACGAATGTGCTGT 373

RESULT 33
US-09-949-016-33748/c
; Sequence 33748, Application US/09949016

```
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 33748
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-33748

Query Match      68.0%; Score 17; DB 3; Length 601;
Best Local Similarity 80.0%; Pred. No. 3.3e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAAACAATGACCAACTGGGCTGT 25
   ||||| ||||| ||||| ||||| |||||
Db 193 AATAACAATGATGAATGGTCTGT 169

RESULT 34
US-09-949-016-133372
; Sequence 133372, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 133372
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-133372

Query Match      68.0%; Score 17; DB 3; Length 601;
Best Local Similarity 80.0%; Pred. No. 3.3e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAAACAATGACCAACTGGGCTGT 25
   ||||| ||||| ||||| ||||| |||||
Db 181 AAAAAAATGACCAATGGGCTGT 205

RESULT 35
US-09-949-016-151005
; Sequence 151005, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 151005
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-151005

Query Match      68.0%; Score 17; DB 3; Length 601;
Best Local Similarity 80.0%; Pred. No. 3.3e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAAACAATGACCAACTGGGCTGT 25
   ||||| ||||| ||||| ||||| |||||
Db 311 AAAAAAATGACCAACTGGGCTAT 287

RESULT 37
US-09-949-016-197059/c
; Sequence 197059, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 197059
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-197059/c

Query Match      68.0%; Score 17; DB 3; Length 601;
Best Local Similarity 80.0%; Pred. No. 3.3e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAAACAATGACCAACTGGGCTGT 25
   ||||| ||||| ||||| ||||| |||||
Db 311 AAAAAAATGACCAACTGGGCTAT 287

RESULT 37
US-09-949-016-197059/c
; Sequence 197059, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 197059
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-197059/c

Query Match      68.0%; Score 17; DB 3; Length 601;
Best Local Similarity 80.0%; Pred. No. 3.3e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAAACAATGACCAACTGGGCTGT 25
   ||||| ||||| ||||| ||||| |||||
Db 311 AAAAAAATGACCAACTGGGCTAT 287
```

Query Match 68.0%; Score 17; DB 3; Length 8888;
Best Local Similarity 100.0%; Pred. No. 4.7e+02;
Matches 17; Conservative 0; Mismatches 0; Indels

```
Query Match      68.0%; Score 17; DB 3; Length 601;
Best Local Similarity 80.0%; Pred. NO. 3.3e+02;
Matches 20: Conservative 0; Mismatches 5; Indels 0; Gaps 0;
```

Qy 1 AAAAAACAATGACCAAC 17
|||
pb 6956 AAAAAACAATGACCAAC 6940

```

1      RESULT 40
2      US-09-949-016-15760/c
3      ; Sequence 15760, Application US/09949016
4      ; Patent No. 6812339
5      ; GENERAL INFORMATION:
6      ; APPLICANT: VENTER, J. Craig et al.
7      ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
8      ; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
9      ; FILE REFERENCE: CL001307
10     ; CURRENT APPLICATION NUMBER: US/09/949,016
11     ; CURRENT FILING DATE: 2000-04-14
12     ; PRIOR APPLICATION NUMBER: 60/241,755
13     ; PRIOR FILING DATE: 2000-10-20
14     ; PRIOR APPLICATION NUMBER: 60/237,768
15     ; PRIOR FILING DATE: 2000-10-03
16     ; PRIOR APPLICATION NUMBER: 60/231,498
17     ; PRIOR FILING DATE: 2000-09-08
18     ; NUMBER OF SEQ ID NOS: 207012
19     ; SOFTWARE: FastSEQ for Windows Version 4.0
20     ; SEQ ID NO 15760
21     ; LENGTH: 8889
22     ; TYPE: DNA
23     ; ORGANISM: Human
24     US-09-949-016-15760

```

```
Query Match      68.0%; Score 17; DB 3; Length 8889;
Best Local Similarity 100.0%; Pred. No. 4.7e+02;
Matches 17: Conservative 0; Mismatches 0 Indels
```

QY 1 AAAAAACAATGACCAAC 17
Db 6956 AAAAAACAATGACCAAC 6940

```

RESULT 41
US-09-949-016-11927
; Sequence 11927, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11927
; LENGTH: 15585
; TYPE: DNA

```

```
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(15585)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-11927

Query Match      68.0%; Score 17; DB 3; Length 15585;
Best Local Similarity 80.0%; Pred. No. 5.1e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 AAAAAACAATGACCAACTGGGCTGT 25
    ||||| || |||| || ||||
Db 10984 AAAAAACAGTGTCACAGGGGACTGT 11008

RESULT 42
US-09-949-016-15627
; Sequence 15627, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15627
; LENGTH: 15585
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(15585)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15627

Query Match      68.0%; Score 17; DB 3; Length 15585;
Best Local Similarity 80.0%; Pred. No. 5.1e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 AAAAAACAATGACCAACTGGGCTGT 25
    ||||| || |||| || ||||
Db 10984 AAAAAACAGTGTCACAGGGGACTGT 11008

RESULT 43
US-09-949-016-15577
; Sequence 15577, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15577
; LENGTH: 15577
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(15577)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15577

Query Match      68.0%; Score 17; DB 3; Length 15577;
Best Local Similarity 80.0%; Pred. No. 5.8e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 AAAAAACAATGACCAACTGGGCTGT 25
    ||||| || |||| || ||||
Db 41349 AATAACACATGATGAAATGGTCTGT 41325

RESULT 44
US-09-949-016-12550/c
; Sequence 12550, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12550
; LENGTH: 45842
; TYPE: DNA
; ORGANISM: Human
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12550

Query Match      68.0%; Score 17; DB 3; Length 45842;
Best Local Similarity 80.0%; Pred. No. 5.8e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 AAAAAACAATGACCAACTGGGCTGT 25
    ||||| || |||| || ||||
Db 41349 AATAACACATGATGAAATGGTCTGT 41325

RESULT 45
US-09-949-016-17327/c
; Sequence 17327, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17327
```

```
; LENGTH: 45842
; TYPE: DNA
; ORGANISM: Human
; NAME/KEY: misc_feature
; LOCATION: (1)...(56241)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17327

Query Match      68.0%; Score 17; DB 3; Length 45842;
Best Local Similarity 80.0%; Pred. No. 5.9e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAAAATGACCAACTGGGCTGT 25
   ||||| ||||| ||||| |||||
Db 41349 AATAACAATGATGAAATGGTCTGT 41325

RESULT 46
US-09-949-016-15174
; Sequence 15174, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15174
; LENGTH: 56241
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(56241)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15174

Query Match      68.0%; Score 17; DB 3; Length 56241;
Best Local Similarity 80.0%; Pred. No. 5.9e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAAAATGACCAACTGGGCTGT 25
   ||||| ||||| ||||| |||||
Db 41244 AAAAAATCATGACCAATGGGCTCT 41268

RESULT 47
US-09-949-016-15175
; Sequence 15175, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15175
; LENGTH: 56241
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(56241)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15175

Query Match      68.0%; Score 17; DB 3; Length 56241;
Best Local Similarity 80.0%; Pred. No. 5.9e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAAAATGACCAACTGGGCTGT 25
   ||||| ||||| ||||| |||||
Db 1516 AATAGCAATGATCAAAATGGGCTCT 1492

RESULT 49
US-09-949-002-799/c
; Sequence 799, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL000790
; CURRENT APPLICATION NUMBER: US/09/949,002
; CURRENT FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/231,401
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 10823
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 799
; LENGTH: 75295
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(75295)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-002-575

Query Match      68.0%; Score 17; DB 3; Length 75295;
Best Local Similarity 80.0%; Pred. No. 6.2e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAAAATGACCAACTGGGCTGT 25
   ||||| ||||| ||||| |||||
Db 1516 AATAGCAATGATCAAAATGGGCTCT 1492

RESULT 48
US-09-949-002-575/c
; Sequence 575, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL000790
; CURRENT APPLICATION NUMBER: US/09/949,002
; CURRENT FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/231,401
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 10823
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 575
; LENGTH: 75295
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(75295)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-002-575

Query Match      68.0%; Score 17; DB 3; Length 75295;
Best Local Similarity 80.0%; Pred. No. 6.2e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAAAATGACCAACTGGGCTGT 25
   ||||| ||||| ||||| |||||
Db 1516 AATAGCAATGATCAAAATGGGCTCT 1492

RESULT 49
US-09-949-002-799/c
; Sequence 799, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL000790
; CURRENT APPLICATION NUMBER: US/09/949,002
; CURRENT FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/231,401
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 10823
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 799
; LENGTH: 75296
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
```

```
; NAME/KEY: misc feature
; LOCATION: (1)_(75296)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-002-799

Query Match      68.0%; Score 17; DB 3; Length 75296;
Best Local Similarity 80.0%; Pred. No. 6.2e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAAACAATGACCAACTGGGCTGT 25
   ||||| ||||| ||||| |||||
Db 1516 AATAGCAATGATCAAAATGGGCTCT 1492

RESULT 50
US-09-949-016-11990
; Sequence 11990, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq For Windows Version 4.0
; SEQ ID NO 11990
; LENGTH: 99748
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-11990

Query Match      68.0%; Score 17; DB 3; Length 99748;
Best Local Similarity 80.0%; Pred. No. 6.3e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAAACAATGACCAACTGGGCTGT 25
   ||||| ||||| ||||| |||||
Db 60414 AAAAAAAAAAACAACCAACTGGGTAT 60438

Search completed: February 3, 2006, 16:32:05
Job time : 88.6667 secs
```

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 3, 2006, 14:37:28 ; Search time 365.556 Seconds
(without alignments)
565.535 Million cell updates/sec

Title: US-10-719-900-5

Perfect score: 25

Sequence: 1 aaaaacaatgacaaactgggtgt 25

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database : Published Applications NA Main:

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
- 3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
- 4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
- 6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
- 8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	25	100.0	25	8	US-10-719-900-5
2	23.4	93.6	25	8	US-10-719-900-6
3	20.8	83.2	598	4	US-09-925-065A-832538
4	20.2	80.8	553	4	US-09-925-065A-334942
5	20.2	80.8	553	4	US-09-925-065A-334943
6	20.2	80.8	585	4	US-09-925-065A-908094
7	19.4	77.6	2316	7	US-10-437-963-67029
8	19.2	76.8	460	3	US-10-437-963-50993
9	19.2	76.8	600	3	US-09-864-761-13549
10	19.2	76.8	1182	7	US-10-437-963-55485
11	18.8	75.2	6596	10	US-11-097-143-9349
12	18.6	74.4	171843	9	US-10-981-277-44
13	18.6	74.4	290547	7	US-10-367-094-77
14	18.4	73.6	308	3	US-09-954-531-4
15	18.4	73.6	308	9	US-10-843-641A-1071
16	18.4	73.6	327	7	US-10-242-535A-12176
17	18.4	73.6	327	7	US-10-085-783A-12176
18	18.4	73.6	414	8	US-10-425-115-49124
19	18.4	73.6	471	7	US-10-242-535A-43265
20	18.4	73.6	471	7	US-10-085-783A-43265
21	18.4	73.6	2146	8	US-10-723-860-4899
22	18.4	73.6	2178	8	US-10-425-115-183781
23	18.2	72.8	584	4	US-09-925-065A-187435

24	18.2	72.8	584	4	US-09-925-065A-187436	Sequence 187436,
25	18.2	72.8	594	4	US-09-925-065A-827160	Sequence 827160,
26	18.2	72.8	607	4	US-09-925-065A-584904	Sequence 584904,
27	18.2	72.8	613	4	US-09-925-065A-309641	Sequence 309641,
28	18.2	72.8	650	7	US-10-021-323-17315	Sequence 17315, A
29	18.2	72.8	1003	7	US-10-282-122A-36969	Sequence 36969, A
30	18.2	72.8	1047	7	US-10-282-122A-37716	Sequence 37716, A
31	18.2	72.8	1286	4	US-09-925-065A-676311	Sequence 676311,
32	18.2	72.8	1286	4	US-09-925-065A-676312	Sequence 676312,
33	18.2	72.8	1286	4	US-09-925-065A-676313	Sequence 676313,
34	18.2	72.8	1286	4	US-09-925-065A-676314	Sequence 676314,
35	18.2	72.8	2319	7	US-10-425-114-3874	Sequence 3874, Ap
36	18.2	72.8	2526	8	US-10-425-115-59884	Sequence 59884, A
37	18.2	72.8	54810	8	US-10-417-375-91	Sequence 91, Appl
38	18.2	72.8	66479	5	US-10-041-856-1	Sequence 1, Appl
39	18.2	72.8	66479	10	US-11-073-203-1	Sequence 1, Appl
40	18.2	72.8	78878	8	US-10-719-993-6863	Sequence 6863, Ap
41	18.2	72.8	608916	9	US-10-461-862-1	Sequence 1, Appl
42	18	72.0	1522	3	US-09-938-842A-3818	Sequence 3818, Ap
43	18	72.0	1522	3	US-09-938-842A-3818	Sequence 3818, Ap
44	17.8	71.2	201	8	US-10-741-600-33597	Sequence 33597, A
45	17.8	71.2	379	8	US-10-425-115-171078	Sequence 171078,
46	17.8	71.2	415	8	US-10-425-115-47284	Sequence 47284, A
47	17.8	71.2	612	4	US-09-925-065A-926511	Sequence 926511,
48	17.8	71.2	2570	7	US-10-437-963-41987	Sequence 41987, A
49	17.8	71.2	7600	6	US-10-723-860-6853	Sequence 6853, Ap
50	17.8	71.2	99998	8	US-10-085-117-34	Sequence 34, Appl
51	17.8	71.2	183334	8	US-10-741-600-17646	Sequence 17646, A
52	17.6	70.4	232	8	US-10-425-115-51038	Sequence 51038, A
53	17.6	70.4	374	4	US-09-925-065A-250403	Sequence 250403,
54	17.6	70.4	397	3	US-09-960-352-835	Sequence 835, App
55	17.6	70.4	459	7	US-10-425-114-21919	Sequence 21919, A
56	17.6	70.4	473	5	US-10-425-114-21977	Sequence 21977, A
57	17.6	70.4	516	5	US-10-027-632-139616	Sequence 139616,
58	17.6	70.4	516	5	US-10-027-632-139617	Sequence 139617,
59	17.6	70.4	516	6	US-10-027-632-139616	Sequence 139616,
60	17.6	70.4	534	4	US-09-925-065A-549550	Sequence 549550,
61	17.6	70.4	555	3	US-09-764-868-534	Sequence 534, App
62	17.6	70.4	568	9	US-10-779-543-18513	Sequence 18513, A
63	17.6	70.4	570	4	US-09-925-065A-465338	Sequence 465338,
64	17.6	70.4	570	4	US-09-925-065A-465339	Sequence 465339,
65	17.6	70.4	577	4	US-09-925-065A-783513	Sequence 783513,
66	17.6	70.4	581	4	US-09-925-065A-116999	Sequence 116999,
67	17.6	70.4	581	4	US-09-925-065A-116999	Sequence 116999,
68	17.6	70.4	587	4	US-09-925-065A-190858	Sequence 190858,
69	17.6	70.4	587	5	US-10-027-632-203935	Sequence 203935,
70	17.6	70.4	587	5	US-10-027-632-203936	Sequence 203936,
71	17.6	70.4	587	6	US-10-027-632-203935	Sequence 203935,
72	17.6	70.4	587	6	US-10-027-632-203936	Sequence 203936,
73	17.6	70.4	592	8	US-10-425-115-83254	Sequence 83254, A
74	17.6	70.4	597	4	US-09-925-065A-866728	Sequence 866728,
75	17.6	70.4	600	9	US-10-972-079-40005	Sequence 40005, A
76	17.6	70.4	614	5	US-10-027-632-291895	Sequence 291895,
77	17.6	70.4	614	6	US-10-027-632-291895	Sequence 291895,
78	17.6	70.4	626	4	US-09-925-065A-544804	Sequence 544804,
79	17.6	70.4	628	4	US-09-925-065A-927904	Sequence 927904,
80	17.6	70.4	635	8	US-10-425-115-151662	Sequence 151662,
81	17.6	70.4	669	5	US-10-106-698-643	Sequence 643, App
82	17.6	70.4	741	4	US-09-925-065A-955248	Sequence 955248,
83	17.6	70.4	743	4	US-09-925-065A-8859	Sequence 8859, Ap
84	17.6	70.4	797	9	US-10-779-543-4087	Sequence 4087, Ap
85	17.6	70.4	893	5	US-10-027-632-171163	Sequence 171163,
86	17.6	70.4	893	5	US-10-027-632-171164	Sequence 171164,
87	17.6	70.4	893	6	US-10-027-632-171163	Sequence 171163,
88	17.6	70.4	893	6	US-10-027-632-171164	Sequence 171164,
89	17.6	70.4	1071	4	US-09-925-065A-553293	Sequence 553293,
90	17.6	70.4	1163	7	US-10-425-114-22701	Sequence 22701, A
91	17.6	70.4	1197	4	US-09-925-065A-690699	Sequence 690699,
92	17.6	70.4	1570	7	US-10-433-802-33	Sequence 33, Appl
93	17.6	70.4	1714	3	US-10-425-115-34282	Sequence 34282, A
94	17.6	70.4	1921	8	US-09-764-868-161	Sequence 161, App
95	17.6	70.4	2691	7	US-10-437-963-50361	Sequence 50361, A
96	17.6	70.4	2789	7	US-10-161-493-29	Sequence 29, Appl

```

; SEQ ID NO 5
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-719-900-5
Query Match      100.0%; Score 25; DB 8; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAAACAATGACCAACTGGGCTGT 25
    |||||
Db 1 AAAAAACAATGACCAACTGGGCTGT 25

RESULT 2
US-10-719-900-6
; Sequence 6, Application US/10719900
; Publication No. US20050026164A1
; GENERAL INFORMATION:
; APPLICANT: Xue Mei Zhou
; TITLE OF INVENTION: Methods of Genetic Analysis of Mouse
; FILE REFERENCE: 3528.1
; CURRENT APPLICATION NUMBER: US/10719,900
; CURRENT FILING DATE: 2003-11-20
; PRIOR APPLICATION NUMBER: 60/427,808
; PRIOR FILING DATE: 2002-11-20
; NUMBER OF SEQ ID NOS: 982914
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 6
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-719-900-6
Query Match      93.6%; Score 23.4; DB 8; Length 25;
Best Local Similarity 96.0%; Pred. No. 0.85;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAAAACAATGACCAACTGGGCTGT 25
    |||||
Db 1 AAAAAACAATGACCAACTGGGCTGT 25

RESULT 3
US-09-925-065A-832538
; Sequence 832538, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 832538
; LENGTH: 598
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-832538
```

```

Sequence 27, Appl
Sequence 474, App
Sequence 21, Appl
Sequence 1597, Ap
Sequence 1271, Ap
Sequence 46, Appl
Sequence 1100, Ap
Sequence 28, Appl
Sequence 17865, A
Sequence 3, Appli
Sequence 6, Appli
Sequence 6, Appli
Sequence 11, Appl
Sequence 17991, A
Sequence 97, Appl
Sequence 5729, Ap
Sequence 420, App
Sequence 386143,
Sequence 103, App
Sequence 470865,
Sequence 861934,
Sequence 162723,
Sequence 162724,
Sequence 162723,
Sequence 162724,
Sequence 55258, A
Sequence 55259, A
Sequence 55260, A
Sequence 55261, A
Sequence 2623, Ap
Sequence 37, Appl
Sequence 2623, Ap
Sequence 4333, Ap
Sequence 229810,
Sequence 392728,
Sequence 392729,
Sequence 141128,
Sequence 141129,
Sequence 134080,
Sequence 134081,
Sequence 134082,
Sequence 196559,
Sequence 196560,
Sequence 196559,
Sequence 196560,
Sequence 407, App
Sequence 58397, A
Sequence 58396, A
Sequence 30346, A
Sequence 1174, Ap
Sequence 382, App

Sequence 27, Appl
Sequence 474, App
Sequence 21, Appl
Sequence 1597, Ap
Sequence 1271, Ap
Sequence 46, Appl
Sequence 1100, Ap
Sequence 28, Appl
Sequence 17865, A
Sequence 3, Appli
Sequence 6, Appli
Sequence 6, Appli
Sequence 11, Appl
Sequence 17991, A
Sequence 97, Appl
Sequence 5729, Ap
Sequence 420, App
Sequence 386143,
Sequence 103, App
Sequence 470865,
Sequence 861934,
Sequence 162723,
Sequence 162724,
Sequence 55258, A
Sequence 55259, A
Sequence 55260, A
Sequence 55261, A
Sequence 2623, Ap
Sequence 37, Appl
Sequence 2623, Ap
Sequence 4333, Ap
Sequence 229810,
Sequence 392728,
Sequence 392729,
Sequence 141128,
Sequence 141129,
Sequence 134080,
Sequence 134081,
Sequence 134082,
Sequence 196559,
Sequence 196560,
Sequence 196559,
Sequence 196560,
Sequence 407, App
Sequence 58397, A
Sequence 58396, A
Sequence 30346, A
Sequence 1174, Ap
Sequence 382, App

Sequence 27, Appl
Sequence 474, App
Sequence 21, Appl
Sequence 1597, Ap
Sequence 1271, Ap
Sequence 46, Appl
Sequence 1100, Ap
Sequence 28, Appl
Sequence 17865, A
Sequence 3, Appli
Sequence 6, Appli
Sequence 6, Appli
Sequence 11, Appl
Sequence 17991, A
Sequence 97, Appl
Sequence 5729, Ap
Sequence 420, App
Sequence 386143,
Sequence 103, App
Sequence 470865,
Sequence 861934,
Sequence 162723,
Sequence 162724,
Sequence 55258, A
Sequence 55259, A
Sequence 55260, A
Sequence 55261, A
Sequence 2623, Ap
Sequence 37, Appl
Sequence 2623, Ap
Sequence 4333, Ap
Sequence 229810,
Sequence 392728,
Sequence 392729,
Sequence 141128,
Sequence 141129,
Sequence 134080,
Sequence 134081,
Sequence 134082,
Sequence 196559,
Sequence 196560,
Sequence 196559,
Sequence 196560,
Sequence 407, App
Sequence 58397, A
Sequence 58396, A
Sequence 30346, A
Sequence 1174, Ap
Sequence 382, App
```

ALIGNMENTS

```

RESULT 1
US-10-719-900-5
; Sequence 5, Application US/10719900
; Publication No. US20050026164A1
; GENERAL INFORMATION:
; APPLICANT: Xue Mei Zhou
; TITLE OF INVENTION: Methods of Genetic Analysis of Mouse
; FILE REFERENCE: 3528.1
; CURRENT APPLICATION NUMBER: US/10719,900
; CURRENT FILING DATE: 2003-11-20
; PRIOR APPLICATION NUMBER: 60/427,808
; PRIOR FILING DATE: 2002-11-20
; NUMBER OF SEQ ID NOS: 982914
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
```



```
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_67926C.1
US-10-437-963-67029

Query Match      77.6%; Score 19.4; DB 7; Length 2316;
Best Local Similarity 95.2%; Pred. No. 1.3e+02;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 AAAACAATGACCAACTGGGCT 23
   ||||| ||||| ||||| ||||| |||||
DB 834 AAAATAATGACCAACTGGGCT 854

RESULT 8
US-10-437-963-50993/c
; Sequence 50993, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 50993
; LENGTH: 460
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_53428C.1
US-10-437-963-50993

Query Match      76.8%; Score 19.2; DB 7; Length 460;
Best Local Similarity 87.5%; Pred. No. 1.2e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAACAATGACCAACTGGGCTGT 25
   ||||| ||||| ||||| ||||| |||||
DB 71 AAAAGTGTGACCAACTGGGCTGT 48

RESULT 9
US-09-864-761-13549
; Sequence 13549, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecmics-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
```

```
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
; SEQ ID NO 13549
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC009178.4
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.3
US-09-864-761-13549

Query Match      76.8%; Score 19.2; DB 3; Length 600;
Best Local Similarity 87.5%; Pred. No. 1.3e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAACAATGACCAACTGGGCTGT 25
   ||||| ||||| ||||| ||||| |||||
DB 370 AAAACCATGACCAACTGGTTGT 393

RESULT 10
US-10-437-963-55485
; Sequence 55485, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 55485
; LENGTH: 1182
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_57489C.1
US-10-437-963-55485
```

Query Match 76.8%; Score 19.2; DB 7; Length 1182;
Best Local Similarity 87.5%; Pred. No. 1.4e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAAAATGACCAACTGGGCTGT 25
||||| ||||||| ||||||| ||
DB 1104 AAAAAAGTGTGACCAACTGGGCTGT 1127

RESULT 11
US-11-097-143-9349/c
; Sequence 9349, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; TITLE OF INVENTION: DROSOPHILA GENES.
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9349
; LENGTH: 6596
; TYPE: DNA
; ORGANISM: DROSOPHILA
US-11-097-143-9349

Query Match 75.2%; Score 18.8; DB 10; Length 6596;
Best Local Similarity 90.9%; Pred. No. 3e+02;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AAAAAAATGACCAACTGGGCT 23
||||| ||||||| ||||||| ||
DB 6184 AAAAAAATGACCAACTGGACT 6163

RESULT 12
US-10-981-277-44
; Sequence 44, Application US/10981277
; Publication No. US20050181389A1
; GENERAL INFORMATION:
; APPLICANT: Harris, Cole
; APPLICANT: Davis, Lisa
; TITLE OF INVENTION: Compositions and Methods for Glioma Classification
; FILE REFERENCE: 03-968-US
; CURRENT APPLICATION NUMBER: US/10/981,277
; CURRENT FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/516,817
; PRIOR FILING DATE: 2003-11-03
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 44
; LENGTH: 171843
; TYPE: DNA

; ORGANISM: Homo sapiens
US-10-981-277-44

Query Match 74.4%; Score 18.6; DB 9; Length 171843;
Best Local Similarity 84.0%; Pred. No. 6.7e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAAATGACCAACTGGGCTGT 25
||||| ||||||| ||||||| ||
DB 134483 AAAAAAATGTCCAACAGGGTTGT 134507

RESULT 13
US-10-367-094-77
; Sequence 77, Application US/10367094
; Publication No. US20040170982A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc Malandro
; TITLE OF INVENTION: Novel Therapeutic Targets in Cancer
; FILE REFERENCE: 529452001500
; CURRENT APPLICATION NUMBER: US/10/367,094
; CURRENT FILING DATE: 2003-02-14
; NUMBER OF SEQ ID NOS: 203
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 77
; LENGTH: 290547
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(290547)
; OTHER INFORMATION: n = A,T,C or G
US-10-367-094-77

Query Match 74.4%; Score 18.6; DB 7; Length 290547;
Best Local Similarity 84.0%; Pred. No. 7.3e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAAATGACCAACTGGGCTGT 25
||||| ||||||| ||||||| ||
DB 26055 AAAAAAATATCAACAACACTGGGCTGT 26079

RESULT 14
US-09-954-531-4/c
; Sequence 4, Application US/09954531
; Patent No. US20020165180A1
; GENERAL INFORMATION:
; APPLICANT: Weaver, Zoe
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Can
; TITLE OF INVENTION: Gene Sets
; FILE REFERENCE: 689290-77
; CURRENT APPLICATION NUMBER: US/09/954,531
; CURRENT FILING DATE: 2002-05-02
; PRIOR APPLICATION NUMBER: US/60/233,133
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,009
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,034
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,509
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: US/60/234,567
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 1392
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 308
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-531-4

```
Query Match          73.6%; Score 18.4; DB 3; Length 308;
Best Local Similarity 95.0%; Pred. No. 2.7e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AAAAAACAATGACCAACTGG 20
Db 21 AAAAAACAATGTCCAACCTGG 2

RESULT 15
US-10-843-641A-1071/c
; Sequence 1071, Application US/10843641A
; Publication No. US20050064454A1
; GENERAL INFORMATION:
; APPLICANT: Avalon Pharmaceuticals, Inc.
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
; FILE REFERENCE: 689290-189
; CURRENT APPLICATION NUMBER: US/10/843,641A
; CURRENT FILING DATE: 2004-05-12
; PRIOR APPLICATION NUMBER: US/09/873,367
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US/09/954,531
; PRIOR FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/09/954,456
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/962,436
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/962,832
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/964,824
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US/09/967,768
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US/09/968,007
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US/09/969,347
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US/09/969,708
; PRIOR FILING DATE: 2001-10-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 847
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1071
; LENGTH: 308
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-843-641A-1071

Query Match          73.6%; Score 18.4; DB 9; Length 308;
Best Local Similarity 95.0%; Pred. No. 2.7e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AAAAAACAATGACCAACTGG 20
Db 21 AAAAAACAATGTCCAACCTGG 2

RESULT 16
US-10-242-535A-12176
; Sequence 12176, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liaw, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2005
; CURRENT APPLICATION NUMBER: US/10/242,535A
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 10/085,783
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
```

```
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 12176
; LENGTH: 327
; TYPE: DNA
; ORGANISM: Human
US-10-242-535A-12176

Query Match          73.6%; Score 18.4; DB 7; Length 327;
Best Local Similarity 95.0%; Pred. No. 2.7e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AAAAAACAATGACCAACTGG 20
Db 146 AAAAAACAATGTCCAACCTGG 165

RESULT 17
US-10-085-783A-12176
; Sequence 12176, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liaw, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2002
; CURRENT APPLICATION NUMBER: US/10/085,783A
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 12176
; LENGTH: 327
; TYPE: DNA
; ORGANISM: Human
US-10-085-783A-12176

Query Match          73.6%; Score 18.4; DB 7; Length 327;
Best Local Similarity 95.0%; Pred. No. 2.7e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AAAAAACAATGACCAACTGG 20
Db 146 AAAAAACAATGTCCAACCTGG 165

RESULT 18
US-10-425-115-49124
; Sequence 49124, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 49124
; LENGTH: 414
; TYPE: DNA
```

```
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_144806C.1
US-10-425-115-49124

Query Match      73.6%; Score 18.4; DB 8; Length 414;
Best Local Similarity 95.0%; Pred. No. 2.8e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAAAACAATGACCAACTGG 20
   |||||
Db 38 AAAAAACAATGTCCCAACTGG 57

RESULT 19
US-10-242-535A-43265
; Sequence 43265, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liew, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2005
; CURRENT APPLICATION NUMBER: US/10/242,535A
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 10/085,783
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 43265
; LENGTH: 471
; TYPE: DNA
; ORGANISM: Human
US-10-242-535A-43265

Query Match      73.6%; Score 18.4; DB 7; Length 471;
Best Local Similarity 95.0%; Pred. No. 2.9e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAAAACAATGACCAACTGG 20
   |||||
Db 164 AAAAAACAATGTCCCAACTGG 183

RESULT 20
US-10-085-783A-43265
; Sequence 43265, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liew, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2002
; CURRENT APPLICATION NUMBER: US/10/085,783A
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 43265
; LENGTH: 471
; TYPE: DNA
; ORGANISM: Human
```

```
US-10-085-783A-43265

Query Match      73.6%; Score 18.4; DB 7; Length 471;
Best Local Similarity 95.0%; Pred. No. 2.9e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAAAACAATGACCAACTGG 20
   |||||
Db 164 AAAAAACAATGTCCCAACTGG 183

RESULT 21
US-10-723-860-4899/c
; Sequence 4899, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Nataeha
; APPLICANT: Ginsburg, Wendy M.
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: Methods for Diagnosis of Soft Tissue Sarcoma, Compositions &
; TITLE OF INVENTION: Methods for Screening for Soft Tissue Sarcoma Modulators
; FILE REFERENCE: 05882.0193.NPUS01
; CURRENT APPLICATION NUMBER: US/10/723,860
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: 60/429,739
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 8393
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4899
; LENGTH: 2146
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-723-860-4899

Query Match      73.6%; Score 18.4; DB 8; Length 2146;
Best Local Similarity 95.0%; Pred. No. 3.8e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAAAACAATGACCAACTGG 20
   |||||
Db 679 AAAAAACAATGTCCCAACTGG 660

RESULT 22
US-10-425-115-183781/c
; Sequence 183781, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(5322)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 183781
; LENGTH: 2178
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(2178)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_99191C.1
US-10-425-115-183781

Query Match      73.6%; Score 18.4; DB 8; Length 2178;
Best Local Similarity 95.0%; Pred. No. 3.8e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 AAAAAACAATGACCAACTGG 20
    |||||
Db 2003 AAAAAACAATACCAACTGG 1984
    |||||

RESULT 23
US-09-925-065A-187435
; Sequence 187435, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 187435
; LENGTH: 584
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-187435

Query Match 72.8%; Score 18.2; DB 4; Length 584;
Best Local Similarity 87.0%; Pred. No. 3.7e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 AAAAAACAATGACCAACTGGCTGT 25
    |||||
Db 492 AATAACAATGACCAACTGGCTTT 514
    |||||

RESULT 24
US-09-925-065A-187436
; Sequence 187436, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 187436
; LENGTH: 584
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-187436

Query Match 72.8%; Score 18.2; DB 4; Length 584;
Best Local Similarity 87.0%; Pred. No. 3.7e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 AAAAAACAATGACCAACTGGCTGT 25
    |||||
Db 492 AATAACAATGACCAACTGGCTTT 514
    |||||

RESULT 25
US-09-925-065A-827160
; Sequence 827160, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 827160
; LENGTH: 594
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-827160

Query Match 72.8%; Score 18.2; DB 4; Length 594;
Best Local Similarity 87.0%; Pred. No. 3.7e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAACAATGACCAACTGGCT 23
    |||||
Db 21 AAAAAAATGACAACTGGACT 43
    |||||

RESULT 26
US-09-925-065A-584904
; Sequence 584904, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 584904
; LENGTH: 607
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-584904
```



```
RESULT 30
US-10-282-122A-39716
; Sequence 39716, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haeelbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR FILING DATE: 2000-03-21
; PRIOR FILING DATE: 2000-03-21
; PRIOR FILING DATE: 2000-05-23
; PRIOR FILING DATE: 2000-05-26
; PRIOR FILING DATE: 2000-05-26
; PRIOR FILING DATE: 2000-09-06
; PRIOR FILING DATE: 2000-09-06
; PRIOR FILING DATE: 2000-09-09
; PRIOR FILING DATE: 2000-10-23
; PRIOR FILING DATE: 2000-10-23
; PRIOR FILING DATE: 2000-11-27
; PRIOR FILING DATE: 2000-12-22
; PRIOR FILING DATE: 2000-12-22
; PRIOR FILING DATE: 2001-02-09
; PRIOR FILING DATE: 2001-02-09
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 39716
; LENGTH: 1047
; TYPE: DNA
; ORGANISM: Salmonella typhi
US-10-282-122A-39716

Query Match 72.8%; Score 18.2; DB 7; Length 1047;
Best Local Similarity 87.0%; Pred. No. 4.1e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 AAAAAAATGACCAACTGGGCTG 24
||||| ||||| ||||| ||||| |||||
Db 435 AAAAAACATTACCAGCTGGGCTG 457

RESULT 31
US-09-925-065A-676311/c
; Sequence 676311, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2001-08-08
```

```
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 676311
; LENGTH: 1286
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-676311

Query Match 72.8%; Score 18.2; DB 4; Length 1286;
Best Local Similarity 87.0%; Pred. No. 4.2e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AAAAAACAATGACCAACTGGGCT 23
||||| ||||| ||||| ||||| |||||
Db 1217 AAAAAACCATGACCAACTGAGGT 1195

RESULT 32
US-09-925-065A-676312/c
; Sequence 676312, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 676312
; LENGTH: 1286
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-676312

Query Match 72.8%; Score 18.2; DB 4; Length 1286;
Best Local Similarity 87.0%; Pred. No. 4.2e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AAAAAACAATGACCAACTGGGCT 23
||||| ||||| ||||| ||||| |||||
Db 1217 AAAAAACCATGACCAACTGAGGT 1195

RESULT 33
US-09-925-065A-676313/c
; Sequence 676313, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
```


; PRIOR APPLICATION NUMBER: US 60/243,096
 ; PRIOR FILING DATE: 2000-10-24
 ; PRIOR APPLICATION NUMBER: US 60/252,147
 ; PRIOR FILING DATE: 2000-11-20
 ; PRIOR APPLICATION NUMBER: US 60/250,092
 ; PRIOR FILING DATE: 2000-11-30
 ; PRIOR APPLICATION NUMBER: US 60/261,766
 ; PRIOR FILING DATE: 2001-01-16
 ; PRIOR APPLICATION NUMBER: US 60/289,846
 ; PRIOR FILING DATE: 2001-05-09
 ; NUMBER OF SEQ ID NOS: 957086
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 676313
 ; LENGTH: 1286
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-925-065A-676313

Query Match 72.8%; Score 18.2; DB 4; Length 1286;
 Best Local Similarity 87.0%; Pred. No. 4.2e+02;
 Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAACAATGACCAACTGGGCT 23
 |||||
 Db 1217 AAAAAACCATGACCAACTGAGGT 1195

RESULT 34
 US-09-925-065A-676314/c
 ; Sequence 676314, Application US/09925065A
 ; Publication No. US20050228172A9
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, David G.
 ; TITLE OF INVENTION: Identification and Mapping of Single
 ; FILE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
 ; FILE REFERENCE: 108827.135
 ; CURRENT APPLICATION NUMBER: US/09/925,065A
 ; CURRENT FILING DATE: 2001-08-08
 ; PRIOR APPLICATION NUMBER: US 60/243,096
 ; PRIOR FILING DATE: 2000-10-24
 ; PRIOR APPLICATION NUMBER: US 60/252,147
 ; PRIOR FILING DATE: 2000-11-20
 ; PRIOR APPLICATION NUMBER: US 60/250,092
 ; PRIOR FILING DATE: 2000-11-30
 ; PRIOR APPLICATION NUMBER: US 60/261,766
 ; PRIOR FILING DATE: 2001-01-16
 ; PRIOR APPLICATION NUMBER: US 60/289,846
 ; PRIOR FILING DATE: 2001-05-09
 ; NUMBER OF SEQ ID NOS: 957086
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 676314
 ; LENGTH: 1286
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-925-065A-676314

Query Match 72.8%; Score 18.2; DB 4; Length 1286;
 Best Local Similarity 87.0%; Pred. No. 4.2e+02;
 Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAACAATGACCAACTGGGCT 23
 |||||
 Db 1217 AAAAAACCATGACCAACTGAGGT 1195

RESULT 35
 US-10-425-114-3974
 ; Sequence 3974, Application US/10425114
 ; Publication No. US2004003488A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Liu, Jingdong
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Kovalic, David K.

; APPLICANT: Screen, Steven E
 ; APPLICANT: Tabaska, Jack E
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53313)B
 ; CURRENT APPLICATION NUMBER: US/10/425,114
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 73128
 ; SEQ ID NO 3974
 ; LENGTH: 2319
 ; TYPE: DNA
 ; ORGANISM: Zea mays
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: 700345691_FLI
 ; US-10-425-114-3974

Query Match 72.8%; Score 18.2; DB 7; Length 2319;
 Best Local Similarity 87.0%; Pred. No. 4.7e+02;
 Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAACAATGACCAACTGGGCT 23
 |||||
 Db 2012 AAAAAACATGACGAATGGCCT 2034

RESULT 36
 US-10-425-115-59884
 ; Sequence 59884, Application US/10425115
 ; Publication No. US20040214272A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa, Thomas J.
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants
 ; FILE REFERENCE: 38-21(53222)B
 ; CURRENT APPLICATION NUMBER: US/10/425,115
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 369326
 ; SEQ ID NO 59884
 ; LENGTH: 2526
 ; TYPE: DNA
 ; ORGANISM: Zea mays
 ; FEATURE:
 ; NAME/KEY: unsure
 ; LOCATION: (1)..(2526)
 ; OTHER INFORMATION: unsure at all n locations
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: MRT4577_154612C.1
 ; US-10-425-115-59884

Query Match 72.8%; Score 18.2; DB 8; Length 2526;
 Best Local Similarity 87.0%; Pred. No. 4.8e+02;
 Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAACAATGACCAACTGGGCT 23
 |||||
 Db 2012 AAAAAACATGACGAATGGCCT 2034

RESULT 37
 US-10-417-375-91/c
 ; Sequence 91, Application US/10417375
 ; Publication No. US20040219528A1
 ; GENERAL INFORMATION:
 ; APPLICANT: David W. Morris
 ; APPLICANT: Marc Malandro
 ; TITLE OF INVENTION: Novel Therapeutic Targets in Cancer
 ; FILE REFERENCE: 529452001600
 ; CURRENT APPLICATION NUMBER: US/10/417,375
 ; CURRENT FILING DATE: 2003-04-15

; NUMBER OF SEQ ID NOS: 176
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 91
; LENGTH: 54810
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(54810)
; OTHER INFORMATION: n = A,T,C or G
US-10-417-375-91

Query Match 72.8%; Score 18.2; DB 8; Length 54810;
Best Local Similarity 87.0%; Pred. No. 8.3e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 AAAACAATGACCAACTGGGCTGT 25
DB 42877 ATAACAATGACCAACTGGGATGT 42855

RESULT 38

US-10-041-856-1/c
; Sequence 1, Application US/10041856
; Publication No. US20020169299A1
; GENERAL INFORMATION:
; APPLICANT: SLAUGENHAUPT, SUSAN
; APPLICANT: GUSELLA, JAMES F.
; TITLE OF INVENTION: GENE FOR IDENTIFYING INDIVIDUALS WITH FAMILIAL
; FILE REFERENCE: 1829-4004US1
; CURRENT APPLICATION NUMBER: US/10/041,856
; PRIOR FILING DATE: 2002-07-08
; PRIOR FILING DATE: 2001-01-06
; NUMBER OF SEQ ID NOS: 88
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 66479
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-041-856-1

Query Match 72.8%; Score 18.2; DB 5; Length 66479;
Best Local Similarity 87.0%; Pred. No. 8.6e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAACAATGACCAACTGGGCT 23
DB 60341 AATAACAATGACCAATGGCCT 60319

RESULT 39

US-11-073-203-1/c
; Sequence 1, Application US/11073203
; Publication No. US20050204409A1
; GENERAL INFORMATION:
; APPLICANT: SLAUGENHAUPT, SUSAN
; APPLICANT: GUSELLA, JAMES F.
; TITLE OF INVENTION: GENE FOR IDENTIFYING INDIVIDUALS WITH FAMILIAL
; FILE REFERENCE: 1829-4004US1
; CURRENT APPLICATION NUMBER: US/11/073,203
; CURRENT FILING DATE: 2005-03-04
; PRIOR APPLICATION NUMBER: US/10/041,856
; PRIOR FILING DATE: 2002-01-07
; PRIOR APPLICATION NUMBER: 60/260,080
; PRIOR FILING DATE: 2001-01-06
; NUMBER OF SEQ ID NOS: 88
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 66479
; TYPE: DNA

; ORGANISM: Homo sapiens
US-11-073-203-1

Query Match 72.8%; Score 18.2; DB 10; Length 66479;
Best Local Similarity 87.0%; Pred. No. 8.6e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAACAATGACCAACTGGGCT 23
DB 60341 AATAACAATGACCAATGGCCT 60319

RESULT 40

US-10-719-993-6863/c
; Sequence 6863, Application US/10719993
; Publication No. US20040265849A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001496
; CURRENT APPLICATION NUMBER: US/10/719,993
; CURRENT FILING DATE: 2003-11-24
; NUMBER OF SEQ ID NOS: 55342
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6863
; LENGTH: 78878
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(78878)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-719-993-6863

Query Match 72.8%; Score 18.2; DB 8; Length 78878;
Best Local Similarity 87.0%; Pred. No. 8.9e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAACAATGACCAACTGGGCT 23
DB 66718 AATAACAATGACCAATGGCCT 66696

RESULT 41

US-10-461-862-1
; Sequence 1, Application US/10461862
; Publication No. US20050090434A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc S. Malandro
; TITLE OF INVENTION: Novel Therapeutic Targets in Cancer
; FILE REFERENCE: 529452001800
; CURRENT APPLICATION NUMBER: US/10/461,862
; CURRENT FILING DATE: 2003-06-13
; NUMBER OF SEQ ID NOS: 184
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 608916
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(608916)
; OTHER INFORMATION: n = A,T,C or G
US-10-461-862-1

Query Match 72.8%; Score 18.2; DB 9; Length 608916;
Best Local Similarity 87.0%; Pred. No. 1.2e+03;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAACAATGACCAACTGGGCTG 24
||| ||||| ||||| ||||| |||||

Db 432617 AACAGCAATGACCAACTGGCTG 432639

RESULT 42

US-09-938-842A-3818/c

; Sequence 3818, Application US/09938842A

; Patent No. US20020160378A1

; GENERAL INFORMATION:

; APPLICANT: Harper, Jeff

; APPLICANT: Kreps, Joel

; APPLICANT: Wang, Xun

; APPLICANT: Zhu, Tong

; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING

; FILE REFERENCE: SAME, AND METHODS OF USE

; CURRENT APPLICATION NUMBER: US/09/938,842A

; PRIOR FILING DATE: 2001-08-24

; PRIOR APPLICATION NUMBER: US 60/227,866

; PRIOR FILING DATE: 2000-08-24

; PRIOR APPLICATION NUMBER: US 60/264,647

; PRIOR FILING DATE: 2001-01-16

; PRIOR APPLICATION NUMBER: US 60/300,111

; PRIOR FILING DATE: 2001-06-22

; NUMBER OF SEQ ID NOS: 5379

; SEQ ID NO 3818

; LENGTH: 1522

; TYPE: DNA

; ORGANISM: Arabidopsis thaliana

US-09-938-842A-3818

Query Match 72.0%; Score 18; DB 3; Length 1522;

Best Local Similarity 100.0%; Pred. No. 5.4e+02;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAAACAATGACCAACT 18

|||||

1386 AAAAAACAATGACCAACT 1369

RESULT 43

US-09-938-842A-3818/c

; Sequence 3818, Application US/09938842A

; Publication No. US20040009476A9

; GENERAL INFORMATION:

; APPLICANT: Harper, Jeff

; APPLICANT: Kreps, Joel

; APPLICANT: Wang, Xun

; APPLICANT: Zhu, Tong

; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING

; FILE REFERENCE: SAME, AND METHODS OF USE

; CURRENT APPLICATION NUMBER: US/09/938,842A

; PRIOR FILING DATE: 2001-08-24

; PRIOR APPLICATION NUMBER: US 60/227,866

; PRIOR FILING DATE: 2000-08-24

; PRIOR APPLICATION NUMBER: US 60/264,647

; PRIOR FILING DATE: 2001-01-16

; PRIOR APPLICATION NUMBER: US 60/300,111

; PRIOR FILING DATE: 2001-06-22

; NUMBER OF SEQ ID NOS: 5379

; SEQ ID NO 3818

; LENGTH: 1522

; TYPE: DNA

; ORGANISM: Arabidopsis thaliana

US-09-938-842A-3818

Query Match 72.0%; Score 18; DB 3; Length 1522;

Best Local Similarity 100.0%; Pred. No. 5.4e+02;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAAACAATGACCAACT 18

|||||

1386 AAAAAACAATGACCAACT 1369

RESULT 44

US-10-741-600-33597/c

; Sequence 33597, Application US/10741600

; Publication No. US20050026169A1

; GENERAL INFORMATION:

; APPLICANT: CARGILL, Michele et al.

; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH

; MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: CL001499

; CURRENT APPLICATION NUMBER: US/10/741,600

; CURRENT FILING DATE: 2003-12-22

; NUMBER OF SEQ ID NOS: 73997

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 33597

; LENGTH: 201

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-741-600-33597

Query Match 71.2%; Score 17.8; DB 8; Length 201;

Best Local Similarity 82.6%; Pred. No. 4.6e+02;

Matches 19; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAACAATGACCAACTGGGCT 23

|||||

112 AAAAAACAACACCAACTGGGCT 90

RESULT 45

US-10-425-115-171078/c

; Sequence 171078, Application US/10425115

; Publication No. US20040214272A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

; FILE REFERENCE: 38-21(53222)B

; CURRENT APPLICATION NUMBER: US/10/425,115

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 369326

; SEQ ID NO 171078

; LENGTH: 379

; TYPE: DNA

; ORGANISM: Zea mays

; FEATURE:

; OTHER INFORMATION: Clone ID: MRT4577_87601C.1

US-10-425-115-171078

Query Match

Best Local Similarity 90.5%; Score 17.8; DB 8; Length 379;

Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 AACCAATGACCAACTGGGCTGT 25

|||||

256 AACCAATGACCAACTGGGCTGT 236

RESULT 46

US-10-425-115-47284

; Sequence 47284, Application US/10425115

; Publication No. US20040214272A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

; FILE REFERENCE: 38-21(53222)B

; CURRENT APPLICATION NUMBER: US/10/425,115

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 369326

; SEQ ID NO 171078

; LENGTH: 379

; TYPE: DNA

; ORGANISM: Zea mays

; FEATURE:

; OTHER INFORMATION: Clone ID: MRT4577_87601C.1

; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 47284
; LENGTH: 415
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_143129C.1
US-10-425-115-47284

Query Match 71.2%; Score 17.8; DB 8; Length 415;
Best Local Similarity 90.5%; Pred. No. 5.3e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AAAAAACAATGACCAACTGGC 22
||||| : |||||
Db 114 AAAAAACAATTAACAACCTGGC 134

RESULT 47
US-09-925-065A-926511/c
; Sequence 926511, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 926511
; LENGTH: 612
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-926511

Query Match 71.2%; Score 17.8; DB 4; Length 612;
Best Local Similarity 82.6%; Pred. No. 5.7e+02;
Matches 19; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAACAATGACCAACTGGCT 23
||||| : |||||
Db 186 AAAAAACAACARCAAACTGGCT 164

RESULT 48
US-10-437-963-41987/c
; Sequence 41987, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 41987
; LENGTH: 2570
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_45280C.1
US-10-437-963-41987

Query Match 71.2%; Score 17.8; DB 7; Length 2570;
Best Local Similarity 90.5%; Pred. No. 7.3e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAACAATGACCAACTGGG 21
||||| : |||||
Db 1093 AAAAAATAATGACCAACTGCG 1073

RESULT 49
US-10-723-860-6853/c
; Sequence 6853, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsburg, Wendy M.
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: Methods for Diagnosis of Soft Tissue Sarcoma, Compositions &
; FILE REFERENCE: 05882.0193.NPUS01
; CURRENT APPLICATION NUMBER: US/10/723,860
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: 60/429,739
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 8393
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6853
; LENGTH: 7600
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (2002)..(2034)
; OTHER INFORMATION: n is a, c, g, or t
US-10-723-860-6853

Query Match 71.2%; Score 17.8; DB 8; Length 7600;
Best Local Similarity 90.5%; Pred. No. 8.9e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 AACAAATGACCAACTGGGCTGT 25
||||| : |||||
Db 4359 AACAAAGGACCCCACTGGGCTGT 4339

RESULT 50
US-10-085-117-34/c
; Sequence 34, Application US/10085117
; Publication No. US2003023234A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
; FILE REFERENCE: 529452000121
; CURRENT APPLICATION NUMBER: US/10/085,117
; CURRENT FILING DATE: 2002-02-27
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 361
; SOFTWARE: FastSeq for Windows Version 4.0

Mon Feb 6 12:23:20 2006

```
; SEQ ID NO 34
; LENGTH: 99998
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: variation
; LOCATION: (1)...(99998)
; OTHER INFORMATION: n = any nucleotide
US-10-085-117-34

Query Match      71.2%; Score 17.8; DB 6; Length 99998;
Best Local Similarity 90.5%; Pred. No. 1.4e+03;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      2 AAAAACAATGACCACTGGGC 22
      ||||| ||||| ||||| |||||
Db      98973 AAAAACAACCACTGGGC 98953

Search completed: February 3, 2006, 15:43:34
Job time : 371.556 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 3, 2006, 11:35:31 ; Search time 343.111 Seconds
(without alignments)
61.064 Million cell updates/sec

Title: US-10-719-900-5

Perfect score: 25

Sequence: 1 aaaaacaatgacaaactgggctgt 25

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 6068529 seqs, 419036697 residues

Total number of hits satisfying chosen parameters: 12137058

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 150 summaries

Database :

- Published Applications NA New:**
- 1: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:
 - 2: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:
 - 3: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:
 - 4: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:
 - 5: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:
 - 6: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:
 - 7: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:
 - 8: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:
 - 9: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:
 - 10: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:
 - 11: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
C 1	17.8	71.2	100000	8	US-11-124-368A-2898 Sequence 2898, Ap
C 2	17.8	71.2	186442	8	US-11-121-086-104 Sequence 104, App
C 3	17.6	70.4	201	7	US-10-995-561-82214 Sequence 82214, A
C 4	17.6	70.4	195235	7	US-10-995-561-13495 Sequence 13495, A
C 5	17.6	70.4	403278	7	US-10-995-561-13421 Sequence 13421, A
C 6	17.4	69.6	83528	7	US-10-995-561-13343 Sequence 13343, A
C 7	17.2	68.8	76589	7	US-10-995-561-13322 Sequence 13322, A
C 8	17	68.0	842	7	US-10-750-185-42851 Sequence 42851, A
C 9	17	68.0	842	7	US-10-750-623-42851 Sequence 42851, A
C 10	17	68.0	1400	8	US-11-136-527-7256 Sequence 7256, Ap
C 11	17	68.0	1400	8	US-11-136-527-8181 Sequence 8181, Ap
C 12	17	68.0	1452	7	US-10-750-185-45050 Sequence 45050, A
C 13	17	68.0	1452	7	US-10-750-623-45050 Sequence 45050, A
C 14	17	68.0	1500	7	US-10-750-185-50014 Sequence 50014, A
C 15	17	68.0	1500	7	US-10-750-623-50014 Sequence 50014, A
C 16	17	68.0	1626	7	US-10-750-185-25035 Sequence 25035, A
C 17	17	68.0	1626	7	US-10-750-623-25035 Sequence 25035, A
C 18	17	68.0	1854	7	US-10-750-185-31957 Sequence 31957, A
C 19	17	68.0	1854	7	US-10-750-623-31957 Sequence 31957, A
C 20	17	68.0	2001	8	US-11-043-752-1189 Sequence 1189, Ap
C 21	17	68.0	2044	8	US-11-136-527-4085 Sequence 4085, Ap
C 22	17	68.0	2479	7	US-10-750-185-42601 Sequence 42601, A

Sequence 42601, A
Sequence 54891, A
Sequence 54891, A
Sequence 3160, Ap
Sequence 2922, Ap
Sequence 62003, A
Sequence 62003, A
Sequence 2914, Ap
Sequence 35994, A
Sequence 35994, A
Sequence 25048, A
Sequence 25048, A
Sequence 2030, Ap
Sequence 13214, A
Sequence 13319, A
Sequence 13444, A
Sequence 1, Appli
Sequence 82, Appli
Sequence 36, Appli
Sequence 13236, A
Sequence 34, Appli
Sequence 15, Appli
Sequence 56, Appli
Sequence 55, Appli
Sequence 211, Appli
Sequence 42575, A
Sequence 72433, A
Sequence 46992, A
Sequence 46992, A
Sequence 43537, A
Sequence 43537, A
Sequence 13162, A
Sequence 13162, A
Sequence 54596, A
Sequence 54596, A
Sequence 30415, A
Sequence 30415, A
Sequence 33041, A
Sequence 33041, A
Sequence 117, App
Sequence 13434, A
Sequence 20, Appli
Sequence 13293, A
Sequence 13293, A
Sequence 13286, A
Sequence 426658, A
Sequence 40865, A
Sequence 41669, A
Sequence 42642, A
Sequence 44331, A
Sequence 50297, A
Sequence 71235, A
Sequence 71259, A
Sequence 13096, A
Sequence 13097, A
Sequence 2807, Ap
Sequence 31949, A
Sequence 31949, A
Sequence 500, App
Sequence 58987, A
Sequence 58987, A
Sequence 49983, A
Sequence 49983, A
Sequence 33465, A
Sequence 33465, A
Sequence 53961, A
Sequence 53961, A
Sequence 130, App
Sequence 39528, A
Sequence 39528, A
Sequence 38349, A
Sequence 38349, A
Sequence 51136, A

```
c 96      16      64.0      2030      7      US-10-750-623-51136      Sequence 51136, A
c 97      16      64.0      2162      7      US-10-750-185-58286      Sequence 58286, A
c 98      16      64.0      2162      7      US-10-750-623-58286      Sequence 58286, A
c 99      16      64.0      2235      7      US-10-750-185-46293      Sequence 46293, A
c 100     16      64.0      2235      7      US-10-750-623-46293      Sequence 46293, A
c 101     16      64.0      2289      7      US-10-750-185-25692      Sequence 25692, A
c 102     16      64.0      2289      7      US-10-750-623-25692      Sequence 25692, A
c 103     16      64.0      2389      7      US-10-750-185-37454      Sequence 37454, A
c 104     16      64.0      2389      7      US-10-750-623-37454      Sequence 37454, A
c 105     16      64.0      2552      7      US-10-750-185-26948      Sequence 26948, A
c 106     16      64.0      2552      7      US-10-750-623-26948      Sequence 26948, A
c 107     16      64.0      3169      7      US-10-750-185-25665      Sequence 25665, A
c 108     16      64.0      3169      7      US-10-750-623-25665      Sequence 25665, A
c 109     16      64.0      32070     8      US-10-995-561-13317      Sequence 13317, A
c 110     16      64.0      32070     8      US-11-124-368A-2907      Sequence 2907, Ap
c 111     16      64.0      150450     8      US-11-112-908-54         Sequence 54, Appl
c 112     16      64.0      165156     7      US-10-995-561-13304      Sequence 13304, A
c 113     16      64.0      175416     8      US-11-121-086-43         Sequence 43, Appl
c 114     16      64.0      175673     8      US-11-121-086-55         Sequence 55, Appl
c 115     16      64.0      191343     8      US-11-112-908-53         Sequence 53, Appl
c 116     16      64.0      212805     8      US-11-112-908-19         Sequence 19, Appl
c 117     16      64.0      403278     7      US-10-995-561-13421      Sequence 13421, A
c 118     16      64.0      1135000     7      US-10-995-561-13286      Sequence 13286, A
c 119     15.8     63.2      600      7      US-10-750-185-782        Sequence 782, App
c 120     15.8     63.2      600      7      US-10-750-623-782        Sequence 782, App
c 121     15.8     63.2      600      8      US-11-136-527-8000        Sequence 8000, Ap
c 122     15.8     63.2      1176     7      US-10-750-185-29230      Sequence 29230, A
c 123     15.8     63.2      1176     7      US-10-750-623-29230      Sequence 29230, A
c 124     15.8     63.2      1260     7      US-10-750-185-52954      Sequence 52954, A
c 125     15.8     63.2      1260     7      US-10-750-623-52954      Sequence 52954, A
c 126     15.8     63.2      1889     7      US-10-750-185-39394      Sequence 39394, A
c 127     15.8     63.2      4431     7      US-10-750-623-39394      Sequence 39394, A
c 128     15.8     63.2      4431     7      US-10-750-185-30584      Sequence 30584, A
c 129     15.8     63.2      4431     7      US-10-750-623-30584      Sequence 30584, A
c 130     15.8     63.2      5415     7      US-10-775-169-153        Sequence 153, App
c 131     15.8     63.2      5580     8      US-11-136-527-3904        Sequence 3904, Ap
c 132     15.6     62.4      22      7      US-10-310-914A-426659      Sequence 426659,
c 133     15.6     62.4      201     7      US-10-995-561-28081      Sequence 28081, A
c 134     15.6     62.4      396     7      US-10-769-744-418        Sequence 418, App
c 135     15.6     62.4      396     8      US-11-096-191-526        Sequence 526, App
c 136     15.6     62.4      529     8      US-11-128-061-2221        Sequence 2221, Ap
c 137     15.6     62.4      529     8      US-11-128-061-5863        Sequence 5863, Ap
c 138     15.6     62.4      529     8      US-11-128-049-2221        Sequence 2221, Ap
c 139     15.6     62.4      529     8      US-11-128-049-5863        Sequence 5863, Ap
c 140     15.6     62.4      529     8      US-10-750-185-4247        Sequence 4247, Ap
c 141     15.6     62.4      600     7      US-10-750-623-4247        Sequence 4247, Ap
c 142     15.6     62.4      600     7      US-10-750-185-50951      Sequence 50951, A
c 143     15.6     62.4      732     7      US-10-750-623-50951      Sequence 50951, A
c 144     15.6     62.4      732     7      US-11-145-703-39         Sequence 39, Appl
c 145     15.6     62.4      985     8      US-10-750-185-55379      Sequence 55379, A
c 146     15.6     62.4      1084     7      US-10-750-623-55379      Sequence 55379, A
c 147     15.6     62.4      1084     7      US-10-750-185-48485      Sequence 48485, A
c 148     15.6     62.4      1116     7      US-10-750-623-48485      Sequence 48485, A
c 149     15.6     62.4      1116     7      US-10-750-185-64025      Sequence 64025, A
c 150     15.6     62.4      1281     7      US-10-750-185-64025      Sequence 64025, A

; PRIOR APPLICATION NUMBER: US 60/568,845
; PRIOR FILING DATE: 2004-05-07
; PRIOR APPLICATION NUMBER: US 60/625,936
; PRIOR FILING DATE: 2004-11-09
; NUMBER OF SEQ ID NOS: 21112
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2898
; LENGTH: 100000
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-124-368A-2898

Query Match      71.2%; Score 17.8; DB 8; Length 100000;
Best Local Similarity 90.5%; Pred. No. 1.4e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy      2      AAAAAAATGACCAACTGGGC 22
      ||||| ||||| ||||| |||||
Db      93082 AAAAAAATGACCAACTGGGC 93062
      ||||| ||||| ||||| |||||

RESULT 2
US-11-121-086-104
; Sequence 104, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; APPLICANT: NIELSEN, KIRSTEN V.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138.6000-00000
; CURRENT APPLICATION NUMBER: US/11/121,086
; CURRENT FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 104
; LENGTH: 186442
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-104

Query Match      71.2%; Score 17.8; DB 8; Length 186442;
Best Local Similarity 90.5%; Pred. No. 1.5e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy      4      AAAAAATGACCAACTGGGCTG 24
      ||||| ||||| ||||| |||||
Db      129637 AAAAAATGACCAACTGGGCTG 129657
      ||||| ||||| ||||| |||||

RESULT 3
US-10-995-561-82214/c
; Sequence 82214, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 82214
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-82214

Query Match      70.4%; Score 17.6; DB 7; Length 201;
Best Local Similarity 83.3%; Pred. No. 63;


```


Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AAACAATGACCAACTGGGCTGT 25
Db 32 ACAACCAATGACCAAGTGGGCTTT 9

RESULT 4

US-10-995-561-13495/c
; Sequence 13495, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13495
; LENGTH: 195235
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-13495

Query Match 70.4%; Score 17.6; DB 7; Length 195235;

Best Local Similarity 83.3%; Pred. No. 1.9e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AAACAATGACCAACTGGGCTGT 25
Db 106040 ACAACCAATGACCAAGTGGGCTTT 106017

RESULT 5

US-10-995-561-13421
; Sequence 13421, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13421
; LENGTH: 403278
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(403278)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-995-561-13421

Query Match 70.4%; Score 17.6; DB 7; Length 403278;

Best Local Similarity 83.3%; Pred. No. 2e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AAAAAAATGACCAACTGGGCTGT 25
Db 18688 AAAAACAATGATGCCAAGGCTGT 18711

RESULT 6

US-10-995-561-13343/c
; Sequence 13343, Application US/10995561
; Publication No. US20050272054A1

; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13343
; LENGTH: 83528
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-13343

Query Match 69.6%; Score 17.4; DB 7; Length 83528;

Best Local Similarity 94.7%; Pred. No. 2e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAAAAATGACCAACTG 19
Db 19575 AAAAAAATGTCCTCACTG 19557

RESULT 7

US-10-995-561-13322
; Sequence 13322, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13322
; LENGTH: 76589
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(76589)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-995-561-13322

Query Match 68.8%; Score 17.2; DB 7; Length 76589;

Best Local Similarity 86.4%; Pred. No. 2.4e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 AAACAATGACCAACTGGGCTGT 25
Db 64823 AATCAATGACCAACTGTGCTGT 64844

RESULT 8

US-10-750-185-42851
; Sequence 42851, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185

```
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 42851
; LENGTH: 842
; TYPE: DNA
; ORGANISM: Bovine 19866881077507
US-10-750-185-42851

Query Match      68.0%; Score 17; DB 7; Length 842;
Best Local Similarity 80.0%; Pred. No. 1.5e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 AAAAAACAATGACCAACTGGGCTGT 25
Db 405 AATAATCAATCACCACACTGCTGT 429

RESULT 9
US-10-750-623-42851
; Sequence 42851, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 42851
; LENGTH: 842
; TYPE: DNA
; ORGANISM: Bovine 19866881077507
US-10-750-623-42851

Query Match      68.0%; Score 17; DB 7; Length 842;
Best Local Similarity 80.0%; Pred. No. 1.5e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 AAAAAACAATGACCAACTGGGCTGT 25
Db 405 AATAATCAATCACCACACTGCTGT 429

RESULT 10
US-11-136-527-7256/c
; Sequence 7256, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 7256
; LENGTH: 1400
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-7256

Query Match      68.0%; Score 17; DB 8; Length 1400;
Best Local Similarity 80.0%; Pred. No. 1.6e+02;
Matches 17; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 1 AAAAAACAATGACCAACTGGGCTGT 25
Db 339 AATAAARATRAYTGACTGGGCTRT 315

RESULT 11
US-11-136-527-8181/c
; Sequence 8181, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 8181
; LENGTH: 1400
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-8181

Query Match      68.0%; Score 17; DB 8; Length 1400;
Best Local Similarity 80.0%; Pred. No. 1.6e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 AAAAAACAATGACCAACTGGGCTGT 25
Db 339 AATAAATAATGATTGACTGGGCTGT 315

RESULT 12
US-10-750-185-45050
; Sequence 45050, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 45050
; LENGTH: 1452
; TYPE: DNA
; ORGANISM: Bovine 19866881030841
US-10-750-185-45050

Query Match      68.0%; Score 17; DB 7; Length 1452;
Best Local Similarity 80.0%; Pred. No. 1.6e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 AAAAAACAATGACCAACTGGGCTGT 25
```

Db 694 AGAAACATCACCATTAGGGCTGT 718
|||||

RESULT 13

US-10-750-623-45050
; Sequence 45050, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 45050
; LENGTH: 1452
; TYPE: DNA
; ORGANISM: Bovine 19866881030841
US-10-750-623-45050

Query Match 68.0%; Score 17; DB 7; Length 1452;
Best Local Similarity 80.0%; Pred. No. 1.6e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAAACAATGACCAACTGGGCTGT 25
|||||

Db 694 AGAAACATCACCATTAGGGCTGT 718
|||||

RESULT 14

US-10-750-185-50014
; Sequence 50014, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 50014
; LENGTH: 1500
; TYPE: DNA
; ORGANISM: Bovine 19866880741238
US-10-750-185-50014

Query Match 68.0%; Score 17; DB 7; Length 1500;
Best Local Similarity 80.0%; Pred. No. 1.6e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAAACAATGACCAACTGGGCTGT 25
|||||

Db 64 AATAAATTATGACCAACTGGGGTTT 88
|||||

RESULT 15

US-10-750-623-50014
; Sequence 50014, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 50014
; LENGTH: 1500
; TYPE: DNA
; ORGANISM: Bovine 19866880741238
US-10-750-623-50014

Query Match 68.0%; Score 17; DB 7; Length 1500;
Best Local Similarity 80.0%; Pred. No. 1.6e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAAACAATGACCAACTGGGCTGT 25
|||||

Db 64 AATAAATTATGACCAACTGGGGTTT 88
|||||

RESULT 16

US-10-750-185-25035/c
; Sequence 25035, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 25035
; LENGTH: 1626
; TYPE: DNA
; ORGANISM: Bovine 19866880851773
US-10-750-185-25035

Query Match 68.0%; Score 17; DB 7; Length 1626;
Best Local Similarity 80.0%; Pred. No. 1.6e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAAACAATGACCAACTGGGCTGT 25
|||||

Db 352 AAAAAAGAAAGAAAACTGGGCAGT 328
|||||

RESULT 17

US-10-750-623-25035/c

; Sequence 25035, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 25035
; LENGTH: 1626
; TYPE: DNA
; ORGANISM: Bovine 19866808051773
US-10-750-623-25035

Query Match 68.0%; Score 17; DB 7; Length 1626;
Best Local Similarity 80.0%; Pred. No. 1.6e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 AAAAAACAATGACCAACTGGGCTGT 25
||||| ||| ||||| |||
Db 352 AAAAAAGAAGAAAACTGGGCAGT 328

RESULT 18
US-10-750-185-31957
; Sequence 31957, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 31957
; LENGTH: 1854
; TYPE: DNA
; ORGANISM: Bovine 198668080691622
US-10-750-185-31957

Query Match 68.0%; Score 17; DB 7; Length 1854;
Best Local Similarity 80.0%; Pred. No. 1.7e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 AAAAAACAATGACCAACTGGGCTGT 25
||||| ||||| ||||| |||
Db 450 AGATAACAATGACCGACTGACCTGT 474

RESULT 19
US-10-750-623-31957
; Sequence 31957, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth

; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 31957
; LENGTH: 1854
; TYPE: DNA
; ORGANISM: Bovine 198668080691622
US-10-750-623-31957

Query Match 68.0%; Score 17; DB 7; Length 1854;
Best Local Similarity 80.0%; Pred. No. 1.7e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 AAAAAACAATGACCAACTGGGCTGT 25
||||| ||||| ||||| |||
Db 450 AGATAACAATGACCGACTGACCTGT 474

RESULT 20
US-11-043-752-1189
; Sequence 1189, Application US/11043752
; Publication No. US20060014165A1
; GENERAL INFORMATION:
; APPLICANT: Hakonarson, Hakon
; APPLICANT: Gurney, Mark E.
; APPLICANT: Halapi, Eva
; TITLE OF INVENTION: METHODS OF DIAGNOSIS AND TREATMENT FOR
; TITLE OF INVENTION: ASTHMA AND OTHER RESPIRATORY DISEASES BASED ON HAPLOTYPE
; FILE REFERENCE: 2345.2044-003
; CURRENT APPLICATION NUMBER: US/11/043,752
; CURRENT FILING DATE: 2005-01-26
; PRIOR APPLICATION NUMBER: PCT/US04/022446
; PRIOR FILING DATE: 2004-07-14
; PRIOR APPLICATION NUMBER: 60/487,072
; PRIOR FILING DATE: 2003-07-14
; PRIOR APPLICATION NUMBER: 60/559,611
; PRIOR FILING DATE: 2004-04-05
; NUMBER OF SEQ ID NOS: 4326
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1189
; LENGTH: 2001
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-043-752-1189

Query Match 68.0%; Score 17; DB 8; Length 2001;
Best Local Similarity 80.0%; Pred. No. 1.7e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 AAAAAACAATGACCAACTGGGCTGT 25
||||| ||||| ||||| |||
Db 336 ATAAACAACAAAACCAACTGGACTAT 360

RESULT 21
US-11-136-527-4085/c
; Sequence 4085, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth

```
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; PRIOR FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4085
; LENGTH: 2044
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-4085

Query Match      68.0%; Score 17; DB 8; Length 2044;
Best Local Similarity 80.0%; Pred. No. 1.7e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAAACAATGACCACTGGGCTGT 25
Db 983 AATAAAATGATTGACTGGGCTGT 959

RESULT 22
US-10-750-185-42601/c
; Sequence 42601, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 42601
; LENGTH: 2479
; TYPE: DNA
; ORGANISM: Bovine 19866880544411
US-10-750-185-42601

Query Match      68.0%; Score 17; DB 7; Length 2479;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAAACAATGACCAAC 17
Db 1816 AAAAAACAATGACCAAC 1800

RESULT 23
US-10-750-623-42601/c
; Sequence 42601, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
```

```
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 42601
; LENGTH: 2479
; TYPE: DNA
; ORGANISM: Bovine 19866880544411
US-10-750-623-42601

Query Match      68.0%; Score 17; DB 7; Length 2479;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAAACAATGACCAAC 17
Db 1816 AAAAAACAATGACCAAC 1800

RESULT 24
US-10-750-185-54891/c
; Sequence 54891, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 54891
; LENGTH: 3109
; TYPE: DNA
; ORGANISM: Bovine 19866880656839
US-10-750-185-54891

Query Match      68.0%; Score 17; DB 7; Length 3109;
Best Local Similarity 80.0%; Pred. No. 1.8e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAAACAATGACCACTGGGCTGT 25
Db 653 AAAACACAATAGACCACTGGTCTGT 629

RESULT 25
US-10-750-623-54891/c
; Sequence 54891, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
```

```
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 54891
; LENGTH: 3109
; TYPE: DNA
; ORGANISM: Bovine 19866880656839
US-10-750-623-54891

Query Match      68.0%; Score 17; DB 7; Length 3109;
Best Local Similarity 80.0%; Pred. No. 1.8e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAAACAATGACCAACTGGGCTGT 25
   ||||| ||||| ||||| ||||| |||||
DB 653 AAAAAACAATAGACAACTGGTCTGT 629

RESULT 26
US-11-136-527-3160/c
; Sequence 3160, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3160
; LENGTH: 3590
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-3160

Query Match      68.0%; Score 17; DB 8; Length 3590;
Best Local Similarity 68.0%; Pred. No. 1.9e+02;
Matches 17; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAACAATGACCAACTGGGCTGT 25
   ||||| ||||| ||||| ||||| |||||
DB 2529 AAATAAARATRAYTGACTGGGCTRT 2505

RESULT 27
US-11-124-368A-2922/c
; Sequence 2922, Application US/11124368A
; Publication No. US20050287559A1
; GENERAL INFORMATION:
; APPLICANT: Michele Cargill
; APPLICANT: James J. Devlin
; APPLICANT: May Luke
; TITLE OF INVENTION: Genetic Polymorphisms Associated with
; FILE REFERENCE: Vascular Diseases, Methods of Detection and Uses Thereof
; FILE REFERENCE: CL001524
; CURRENT APPLICATION NUMBER: US/11/124,368A
; CURRENT FILING DATE: 2005-05-09
; PRIOR APPLICATION NUMBER: US 60/568,845
; PRIOR FILING DATE: 2004-05-07
; PRIOR APPLICATION NUMBER: US 60/625,936
; PRIOR FILING DATE: 2004-11-09
; NUMBER OF SEQ ID NOS: 21112
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2922
; LENGTH: 100000
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-124-368A-2922
```

```
Query Match      68.0%; Score 17; DB 8; Length 100000;
Best Local Similarity 80.0%; Pred. No. 3.1e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAAACAATGACCAACTGGGCTGT 25
   ||||| ||||| ||||| ||||| |||||
DB 60543 AAAAAACATTTTCCAACTGTGCAGT 60519

RESULT 28
US-10-750-185-62003/c
; Sequence 62003, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 62003
; LENGTH: 2396
; TYPE: DNA
; ORGANISM: Bovine 19866881184825
US-10-750-185-62003

Query Match      67.2%; Score 16.8; DB 7; Length 2396;
Best Local Similarity 90.0%; Pred. No. 2.1e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 AACAATGACCAACTGGGCT 23
   ||||| ||||| ||||| |||||
DB 147 AAAAAATGACCAGCTGGGCT 128

RESULT 29
US-10-750-623-62003/c
; Sequence 62003, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 62003
; LENGTH: 2396
; TYPE: DNA
; ORGANISM: Bovine 19866881184825
US-10-750-623-62003

Query Match      67.2%; Score 16.8; DB 7; Length 2396;
Best Local Similarity 90.0%; Pred. No. 2.1e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 4 AAACAATGACCAACTGGGCT 23
Db 147 AAAAAATGACCACTGGGCT 128

RESULT 30
US-11-124-368A-2914
; Sequence 2914, Application US/111124368A
; Publication No. US20050287559A1
; GENERAL INFORMATION:
; APPLICANT: Michele Cargill
; APPLICANT: James J. Devlin
; APPLICANT: May Luke
; TITLE OF INVENTION: Genetic Polymorphisms Associated with
; TITLE OF INVENTION: Vascular Diseases, Methods of Detection and Uses Thereof
; FILE REFERENCE: CL001524
; CURRENT APPLICATION NUMBER: US/11/124,368A
; PRIOR FILING DATE: 2005-05-09
; PRIOR APPLICATION NUMBER: US 60/568,845
; PRIOR FILING DATE: 2004-05-07
; PRIOR APPLICATION NUMBER: US 60/625,936
; PRIOR FILING DATE: 2004-11-09
; NUMBER OF SEQ ID NOS: 21112
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2914
; LENGTH: 73404
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-124-368A-2914

Query Match 67.2%; Score 16.8; DB 8; Length 73404;
Best Local Similarity 90.8%; Pred. No. 3.7e+02; Indels 0; Gaps 0;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 AAACAATGACCAACTGGGC 22
Db 22412 AAATAATGACCACTGGGC 22431

RESULT 31
US-10-750-185-35994/c
; Sequence 35994, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 35994
; LENGTH: 1601
; TYPE: DNA
; ORGANISM: Bovine 19866881606127
US-10-750-185-35994

Query Match 66.4%; Score 16.6; DB 7; Length 1601;
Best Local Similarity 82.6%; Pred. No. 2.5e+02; Indels 0; Gaps 0;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AAACAATGACCAACTGGGCTG 24
Db 1402 AAAACATCAGTCACCTGGGCTG 1380

RESULT 32
US-10-750-623-35994/c
; Sequence 35994, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 35994
; LENGTH: 1601
; TYPE: DNA
; ORGANISM: Bovine 19866881606127
US-10-750-623-35994

Query Match 66.4%; Score 16.6; DB 7; Length 1601;
Best Local Similarity 82.6%; Pred. No. 2.5e+02; Indels 0; Gaps 0;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AAACAATGACCAACTGGGCTG 24
Db 1402 AAAACATCAGTCACCTGGGCTG 1380

RESULT 33
US-10-750-185-25048
; Sequence 25048, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 25048
; LENGTH: 3558
; TYPE: DNA
; ORGANISM: Bovine 19866880528126
US-10-750-185-25048

Query Match 66.4%; Score 16.6; DB 7; Length 3558;
Best Local Similarity 82.6%; Pred. No. 2.8e+02; Indels 0; Gaps 0;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAAATGACCAACTGGGCT 23
Db 2007 AAAAATGACCAACTGGTT 2029

RESULT 34
```

```
US-10-750-623-25048
; Sequence 25048, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 25048
; LENGTH: 3558
; TYPE: DNA
; ORGANISM: Bovine 19866880528126
US-10-750-623-25048
Query Match 66.4%; Score 16.6; DB 7; Length 3558;
Best Local Similarity 82.6%; Pred. No. 2.8e+02;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAACAATGACCAACTGGGCT 23
||||| ||||| ||||| ||||| |||||

Db 2007 AAAAAACACTGACAACTGTGTT 2029

RESULT 35
US-11-136-527-2030/c
; Sequence 2030, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2030
; LENGTH: 4095
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-2030
Query Match 66.4%; Score 16.6; DB 8; Length 4095;
Best Local Similarity 82.6%; Pred. No. 2.9e+02;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAACAATGACCAACTGGGCT 23
||||| ||||| ||||| ||||| |||||

Db 3110 AAAAAACAACGTCAAACTGGGCT 3088

RESULT 36
US-10-995-561-13214/c
; Sequence 13214, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13214
; LENGTH: 19277
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-13214
Query Match 66.4%; Score 16.6; DB 7; Length 19277;
Best Local Similarity 82.6%; Pred. No. 3.7e+02;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AAAAAACAATGACCAACTGGGCTG 24
||||| ||||| ||||| ||||| |||||

Db 13487 AAAAAACATGACTCACTTGGCTG 13465

RESULT 37
US-10-995-561-13319/c
; Sequence 13319, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13319
; LENGTH: 27509
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-13319
Query Match 66.4%; Score 16.6; DB 7; Length 27509;
Best Local Similarity 82.6%; Pred. No. 3.9e+02;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAACAATGACCAACTGGGCT 23
||||| ||||| ||||| ||||| |||||

Db 15454 AAAAAAAAACCACTGGGCT 15432

RESULT 38
US-10-995-561-13444
; Sequence 13444, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13444
; LENGTH: 29618
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-13444
Query Match 66.4%; Score 16.6; DB 7; Length 29618;
Best Local Similarity 82.6%; Pred. No. 3.9e+02;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
```


QY 2 AAAAAAATGACCAACTGGGCTG 24
|||||
Db 25604 AAAAAACATGACTCACTGGCTG 25626

RESULT 39
US-10-857-780-1/c
; Sequence 1, Application US/10857780
; Publication No. US20050272043A1
; GENERAL INFORMATION:
; APPLICANT: ROTH, RICHARD B.
; APPLICANT: BRAUN, ANDREAS
; APPLICANT: KAMMERER, STEFAN M.
; APPLICANT: NELSON, MATTHEW ROBERTS
; APPLICANT: RENELAND, RIKARD HENRY
; APPLICANT: HOYAL-WRIGHTSON, CAROLYN R.
; TITLE OF INVENTION: METHODS FOR IDENTIFYING RISK OF BREAST CANCER AND TREATMENTS
; FILE OF INVENTION: THEREOF
; FILE REFERENCE: SEQ-4069-CP
; CURRENT APPLICATION NUMBER: US/10/857,780
; CURRENT FILING DATE: 2004-05-28
; PRIOR APPLICATION NUMBER: 10/723,681
; PRIOR FILING DATE: 2003-11-25
; PRIOR APPLICATION NUMBER: 60/490,234
; PRIOR FILING DATE: 2003-07-24
; PRIOR APPLICATION NUMBER: 60/525,239
; PRIOR FILING DATE: 2003-11-25
; NUMBER OF SEQ ID NOS: 4962
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 92600
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (22393)..(22394)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (24675)..(24676)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (31942)..(31942)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (32965)..(32966)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (33724)..(33724)
; OTHER INFORMATION: n is a, c, g, or t
US-10-857-780-1

Query Match 66.4%; Score 16.6; DB 7; Length 92600;
Best Local Similarity 82.6%; Pred. No. 4.6e+02;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAAATGACCAACTGGGCT 23
|||||
Db 32635 AAAAAAATGACCACTGGGCT 32613

RESULT 40
US-10-775-169-82
; Sequence 82, Application US/10775169
; Publication No. US20050287532A9
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Burczynski, Michael
; APPLICANT: Twine, Natalie
; APPLICANT: Dörner, Andrew

; APPLICANT: Trepicchio, William
; TITLE OF INVENTION: Method for Monitoring Drug Activities In Vivo
; FILE REFERENCE: AM101080 (031896-013000)
; CURRENT APPLICATION NUMBER: US/10/775,169
; CURRENT FILING DATE: 2004-02-11
; NUMBER OF SEQ ID NOS: 5278
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 82
; LENGTH: 127917
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-775-169-82

Query Match 66.4%; Score 16.6; DB 7; Length 127917;
Best Local Similarity 82.6%; Pred. No. 4.8e+02;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAAATGACCACTGGGCT 23
|||||
Db 90851 AAAAAAGTATAACCACTGGGCT 90873

RESULT 41
US-11-121-086-36
; Sequence 36, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; APPLICANT: NIELSEN, KIRSTEN V.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138.6000-00000
; CURRENT APPLICATION NUMBER: US/11/121,086
; CURRENT FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 36
; LENGTH: 172649
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-36

Query Match 66.4%; Score 16.6; DB 8; Length 172649;
Best Local Similarity 82.6%; Pred. No. 5e+02;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAAATGACCACTGGGCT 23
|||||
Db 12905 AAAAAATTAATTAACCTGGGCT 12927

RESULT 42
US-10-995-561-13236
; Sequence 13236, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: GL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13236
; LENGTH: 305312
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(305312)

; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-4)
US-10-995-561-13236

Query Match 66.4%; Score 16.6; DB 7; Length 305312;
Best Local Similarity 82.6%; Pred. No. 5.2e+02;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AAAAAAATGACCAACTGGGCTG 24
||||| ||||| ||||| ||||| |||||
Db 216461 AAAAAACCATGACTCCTTGGCTG 216483

RESULT 43
US-10-999-208-15/c
; Sequence 15, Application US/10999208
; Publication No. US20060008448A1
; GENERAL INFORMATION:
; APPLICANT: XU, MINZHEN
; APPLICANT: HUMPHREYS, ROBERT
; TITLE OF INVENTION: INHIBITION OF LI EXPRESSION IN MAMMALIAN CELLS
; FILE REFERENCE: REH-2016US01
; CURRENT APPLICATION NUMBER: US/10/999,208
; PRIOR FILING DATE: 2004-11-29
; PRIOR APPLICATION NUMBER: 10/127,347
; PRIOR FILING DATE: 2002-04-22
; PRIOR APPLICATION NUMBER: 10/054,387
; PRIOR FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: 09/205,995
; PRIOR FILING DATE: 1998-12-04
; PRIOR APPLICATION NUMBER: 09/036,746
; PRIOR FILING DATE: 1998-03-09
; PRIOR APPLICATION NUMBER: 08/661,627
; PRIOR FILING DATE: 1996-06-11
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 15
; LENGTH: 57
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: siRNA construct
US-10-999-208-15

Query Match 65.6%; Score 16.4; DB 6; Length 57;
Best Local Similarity 94.4%; Pred. No. 1.8e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AAAAAAATGACCAACTG 19
||||| ||||| ||||| ||||| |||||
Db 57 AAAAAAATGAGCAACTG 40

RESULT 44
US-11-121-438-34
; Sequence 34, Application US/11121438
; Publication No. US20060014173A1
; GENERAL INFORMATION:
; APPLICANT: Huang, Shi
; TITLE OF INVENTION: PR/SET- Domain Containing Nucleic Acids, Antibodies and Methods of Use
; FILE REFERENCE: P-LJ 5301
; CURRENT APPLICATION NUMBER: US/11/121,438
; PRIOR FILING DATE: 2005-05-03
; PRIOR APPLICATION NUMBER: US/10/200,012
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: US 09/910,478
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 34
; LENGTH: 1490
; TYPE: DNA

; ORGANISM: Human
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (64)...(1473)
US-11-121-438-34
Query Match 65.6%; Score 16.4; DB 8; Length 1490;
Best Local Similarity 94.4%; Pred. No. 3e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 AACCAATGACCAACTGGGC 22
||||| ||||| ||||| ||||| |||||
Db 1165 AACCAATGCCCAACTGGGC 1182

RESULT 45
US-11-112-908-56/c
; Sequence 56, Application US/11112908
; Publication No. US20050260659A1
; GENERAL INFORMATION:
; APPLICANT: Harris, Cole
; APPLICANT: Davis, Lisa M.
; TITLE OF INVENTION: Breast Cancer Biomarkers
; FILE REFERENCE: 04-164-US
; CURRENT APPLICATION NUMBER: US/11/112,908
; CURRENT FILING DATE: 2005-04-22
; PRIOR APPLICATION NUMBER: US 60/564,758
; PRIOR FILING DATE: 2004-04-23
; PRIOR APPLICATION NUMBER: US 60/575,978
; PRIOR FILING DATE: 2004-06-01
; PRIOR APPLICATION NUMBER: US 60/631,702
; PRIOR FILING DATE: 2004-11-30
; PRIOR APPLICATION NUMBER: US 60/633,826
; PRIOR FILING DATE: 2004-12-07
; NUMBER OF SEQ ID NOS: 511
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 56
; LENGTH: 150468
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-112-908-56

Query Match 65.6%; Score 16.4; DB 8; Length 150468;
Best Local Similarity 94.4%; Pred. No. 6e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAAAAATGACCAACT 18
||||| ||||| ||||| ||||| |||||
Db 3315 AAAAAAATGACCAACT 3298

RESULT 46
US-11-112-908-55/c
; Sequence 55, Application US/11112908
; Publication No. US20050260659A1
; GENERAL INFORMATION:
; APPLICANT: Harris, Cole
; APPLICANT: Davis, Lisa M.
; TITLE OF INVENTION: Breast Cancer Biomarkers
; FILE REFERENCE: 04-164-US
; CURRENT APPLICATION NUMBER: US/11/112,908
; CURRENT FILING DATE: 2005-04-22
; PRIOR APPLICATION NUMBER: US 60/564,758
; PRIOR FILING DATE: 2004-04-23
; PRIOR APPLICATION NUMBER: US 60/575,978
; PRIOR FILING DATE: 2004-06-01
; PRIOR APPLICATION NUMBER: US 60/631,702
; PRIOR FILING DATE: 2004-11-30
; PRIOR APPLICATION NUMBER: US 60/633,826
; PRIOR FILING DATE: 2004-12-07
; NUMBER OF SEQ ID NOS: 511
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 55

```
; LENGTH: 193789
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-112-908-55

Query Match      65.6%; Score 16.4; DB 8; Length 193789;
Best Local Similarity 94.4%; Pred. No. 6.1e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAAAACAATGACCAACT 18
Db 53500 AAAAAACAATGACCAACT 53483

RESULT 47
US-11-117-187-211
; Sequence 211, Application US/11117187
; Publication No. US2005026560A1
; GENERAL INFORMATION:
; APPLICANT: PREUSS, DAPHNE
; APPLICANT: COPENHAVER, GREGORY
; TITLE OF INVENTION: PLANT ARTIFICIAL CHROMOSOME COMPOSITIONS AND METHODS
; FILE REFERENCE: ARCD:309US
; CURRENT APPLICATION NUMBER: US/11/117,187
; CURRENT FILING DATE: 2005-04-28
; PRIOR APPLICATION NUMBER: US/09/531,120
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/125,219
; PRIOR FILING DATE: 1999-03-18
; NUMBER OF SEQ ID NOS: 212
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 211
; LENGTH: 1082144
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-11-117-187-211

Query Match      65.6%; Score 16.4; DB 8; Length 1082144;
Best Local Similarity 94.4%; Pred. No. 5.2e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAAAACAATGACCAACT 18
Db 824756 AAAAAACAATGACCAACT 824773

RESULT 48
US-10-995-561-42575
; Sequence 42575, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 42575
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-42575

Query Match      64.8%; Score 16.2; DB 7; Length 201;
Best Local Similarity 85.7%; Pred. No. 2.7e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAAACAATGACCAACTGGGC 22
Db 170 AAAAAACAATGACCAACTGGGC 190

us-10-719-900-5.rnpbn

RESULT 49
US-10-995-561-72433
; Sequence 72433, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 72433
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-72433

Query Match      64.8%; Score 16.2; DB 7; Length 201;
Best Local Similarity 85.7%; Pred. No. 2.7e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAAACAATGACCAACTGGGC 22
Db 170 AAAAAACAATGACCAACTGGGC 190

RESULT 50
US-10-750-185-46992
; Sequence 46992, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 46992
; LENGTH: 1030
; TYPE: DNA
; ORGANISM: Bovine 19866880867784
US-10-750-185-46992

Query Match      64.8%; Score 16.2; DB 7; Length 1030;
Best Local Similarity 85.7%; Pred. No. 3.5e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAACAATGACCAACTGGG 21
Db 813 AAAAAACAATGACCAACTGGG 833

Search completed: February 3, 2006, 16:19:52
Job time : 348.111 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 3, 2006, 10:53:23 ; Search time 863 Seconds
(without alignments)
1646.682 Million cell updates/sec

Title: US-10-719-900-10
Perfect score: 25
Sequence: 1 aaaaacacttggaaggagactca 25

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues
Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Maximum Match 0%
Maximum Match 100%
Listing first 150 summaries

Database : GenEmbl.*
1: gb_ba.*
2: gb_in.*
3: gb_env.*
4: gb_ov.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pr.*
9: gb_ro.*
10: gb_ats.*
11: gb_sy.*
12: gb_un.*
13: gb_vi.*
14: gb_htg.*
15: gb_pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	21.4	85.6	110000	14	AC152182_1	Continuation (2 of
2	21.4	85.6	110000	14	AC155165_2	Continuation (3 of
3	21.4	85.6	200774	9	AC116324	AC116324 Mus muscu
4	20.8	83.2	844	9	BC048869	BC048869 Mus muscu
5	20.8	83.2	859	9	BC024624	BC024624 Mus muscu
6	20.8	83.2	870	9	MUSEGFBPM	M17962 Mouse major
7	20.8	83.2	878	9	BC026378	BC026378 Mus muscu
8	20.2	80.8	412	5	AY493302	AY493302 Anas plat
9	20.2	80.8	158857	14	AC165233	AC165233 Mus muscu
10	20.2	80.8	168803	8	AC142350	AC142350 Pan trogl
11	20.2	80.8	189158	8	AC147666	AC147666 Mus muscu
12	20.2	80.8	196620	14	AC145502	AC145502 Canis fam
13	20.2	80.8	210317	9	AC102363	AC102363 Mus muscu
14	20.2	80.8	214385	14	AC153790	AC153790 Mus muscu
15	20.2	80.8	215892	14	AC152260	AC152260 Bos tauru
16	20.2	80.8	230954	14	AC116092	AC116092 Rattus no
17	19.8	79.2	121010	8	AC140677	AC140677 Homo sapi
18	19.8	79.2	163805	8	AC114959	AC114959 Homo sapi

c 92	18.8	75.2	215728	14	AC099112	AC099112 Rattus no
c 93	18.8	75.2	225152	9	AC132454	AC132454 Mus muscu
c 94	18.8	75.2	237165	9	AC112948	AC112948 Mus muscu
c 95	18.8	75.2	243674	14	AC098471	AC098471 Rattus no
c 96	18.8	75.2	249781	14	AC098954	AC098954 Rattus no
c 97	18.8	75.2	294415	14	AC110321	AC110321 Rattus no
c 98	18.6	74.4	520	10	BV279664	BV279664 S232P644R
c 99	18.6	74.4	558	10	BV401108	BV401108 S229P6164
c 100	18.6	74.4	770	10	BV598075	BV598075 S217P6671
c 101	18.6	74.4	1139	9	MUSLIT221	M12819 Mouse t-cel
c 102	18.6	74.4	1268	6	A95291	A95291 Sequence 22
c 103	18.6	74.4	2424	5	BC075293	BC075293 Xenopus t
c 104	18.6	74.4	5310	9	MLVYT22	Y00157 Murine Lys-
c 105	18.6	74.4	6316	9	MUSMHLT2	M22064 Mouse MHC c
c 106	18.6	74.4	9904	6	CQ728820	CQ728820 Sequence
c 107	18.6	74.4	10327	8	AB002309	AB002309 Human mRN
c 108	18.6	74.4	10343	8	U17195	U17195 Homo sapien
c 109	18.6	74.4	45154	14	AC164871	AC164871 Sorex ara
c 110	18.6	74.4	61004	8	AL357566	AL357566 Human DNA
c 111	18.6	74.4	64849	5	AC098643	AC098643 Takifugu
c 112	18.6	74.4	68166	14	AC084712	AC084712 Homo sapi
c 113	18.6	74.4	71266	14	AP007726	AP007726 Lotus cor
c 114	18.6	74.4	71455	14	AC146319	AC146319 Takifugu
c 115	18.6	74.4	79585	15	AP004908	AP004908 Lotus cor
c 116	18.6	74.4	82859	14	AP007748	AP007748 Lotus cor
c 117	18.6	74.4	83607	14	AC101668	AC101668 Mus muscu
c 118	18.6	74.4	86050	8	AL359647	AL359647 Human DNA
c 119	18.6	74.4	92863	11	GGBLUCUS	AL023516 Gallus ga
c 120	18.6	74.4	103019	5	AC091292	AC091292 Takifugu
c 121	18.6	74.4	105245	14	AP007325	AP007325 Lotus cor
c 122	18.6	74.4	107745	8	AL139042	AL139042 Human DNA
c 123	18.6	74.4	108535	14	AC150835	AC150835 Takifugu
c 124	18.6	74.4	110000	14	AC092450	AC092450 Homo sapi
c 125	18.6	74.4	110000	14	AC095357_1	Continuation (2 of
c 126	18.6	74.4	110000	14	CT010578_1	Continuation (2 of
c 127	18.6	74.4	113027	8	AC090669	AC090669 Homo sapi
c 128	18.6	74.4	115040	15	AC124954	AC124954 Medicago
c 129	18.6	74.4	120302	15	AC152403	AC152403 Medicago
c 130	18.6	74.4	121466	8	AC105310	AC105310 Homo sapi
c 131	18.6	74.4	121479	14	AC160837	AC160837 Medicago
c 132	18.6	74.4	125483	14	AC144888	AC144888 Bos tauru
c 133	18.6	74.4	127360	14	AC099409	AC099409 Felis cat
c 134	18.6	74.4	131264	14	AC104198	AC104198 Mus muscu
c 135	18.6	74.4	134569	8	AC105123	AC105123 Homo sapi
c 136	18.6	74.4	135855	14	AC098701	AC098701 Felis cat
c 137	18.6	74.4	137554	14	AC127618	AC127618 Rattus no
c 138	18.6	74.4	139062	15	AC133863	AC133863 Medicago
c 139	18.6	74.4	144439	15	AC147008	AC147008 Medicago
c 140	18.6	74.4	147908	5	AC091293	AC091293 Takifugu
c 141	18.6	74.4	150539	14	CR363639	CR363639 Danio rer
c 142	18.6	74.4	157145	5	BX293547	BX293547 Zebrafish
c 143	18.6	74.4	157842	8	AC092825	AC092825 Homo sapi
c 144	18.6	74.4	157981	14	AC144693	AC144693 Sus scrof
c 145	18.6	74.4	159314	14	AC134321	AC134321 Felis cat
c 146	18.6	74.4	159793	14	AC120687	AC120687 Rattus no
c 147	18.6	74.4	160359	8	AC092722	AC092722 Homo sapi
c 148	18.6	74.4	161781	14	CR376800	CR376800 Danio rer
c 149	18.6	74.4	162907	14	AC026074	AC026074 Homo sapi
c 150	18.6	74.4	164612	9	AC124697	AC124697 Mus muscu

ALIGNMENTS

RESULT 1	AC152182_1	Sequence split into 5 fragments	LOCUS AC152182	Accession	AC152182
WPCOMMENT	Fragment Name	Begin	End		
	AC152182_0	1	110000		
	AC152182_1	100001	210000		
	AC152182_2	200001	310000		
	AC152182_3	300001	410000		
Continuation (2 of 5) of AC152182 from base 100001 (AC152182 Mus musculus chromosome 16					
400001 452855					
Query Match 85.6%; Score 21.4; DB 14; Length 110000;					
Best Local Similarity 95.7%; Pred. No. 16; Length 110000;					
Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;					
QY	2	AAAAACACTTGGAGGAGACTC 24			
Db	29883	AAAAACTCTTGGAGGAGACTC 29905			
RESULT 2					
AC155165_2/c					
WPCOMMENT					
Sequence split into 6 fragments LOCUS AC155165 Accession AC155165					
Fragment Name Begin End					
AC155165_0 1 110000					
AC155165_1 100001 210000					
AC155165_2 200001 310000					
AC155165_3 300001 410000					
AC155165_4 400001 510000					
AC155165_5 500001 551847					
Continuation (3 of 6) of AC155165 from base 200001 (AC155165 Mus musculus chromosome 16					
85.6%; Score 21.4; DB 14; Length 110000;					
Query Match 85.6%; Score 21.4; DB 14; Length 110000;					
Best Local Similarity 95.7%; Pred. No. 16;					
Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;					
QY	2	AAAAACACTTGGAGGAGACTC 24			
Db	59362	AAAAACTCTTGGAGGAGACTC 59340			
RESULT 3					
AC116324/c					
LOCUS					
AC116324 Mus musculus BAC clone RP23-282P1 from 16, complete sequence.					
DEFINITION					
AC116324					
ACCESSION					
AC116324.6 GI:29244826					
VERSION					
HTG.					
KEYWORDS					
Mus musculus (house mouse)					
SOURCE					
Mus musculus					
ORGANISM					
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;					
Sciurognathi; Muridae; Murinae; Mus.					
REFERENCE					
AUTHORS					
Goyea,E., Haglund,K. and Dignan,G.					
TITLE					
The sequence of Mus musculus BAC clone RP23-282P1					
JOURNAL					
Unpublished (2001)					
REFERENCE					
AUTHORS					
Wilson,R.					
TITLE					
Sequencing of Mus musculus					
JOURNAL					
Unpublished (2001)					
REFERENCE					
AUTHORS					
McPherson,J.D. and Waterston,R.H.					
TITLE					
Direct Submission					
JOURNAL					
Submitted (26-MAR-2003) Genome Sequencing Center, 4444 Forest Park					
REFERENCE					
AUTHORS					
Wilson,R.					
TITLE					
Direct Submission					
JOURNAL					
Submitted (08-NOV-2003) Department of Genetics, Washington					

COMMENT

University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 On Mar 26, 2003 this sequence version replaced gi:24943072.
 ----- Genome Center
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: <http://genome.wustl.edu>
 Contact: submissions@watson.wustl.edu
 ----- Summary Statistics
 Center project name: M_BA0282P01

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. Wes Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu>

SOURCE INFORMATION:

The RPCI-23 BAC Library has been constructed by Kazutoyo Osegawa and Minako Tateno in the laboratory of Pieter de Jong (<http://www.chori.org>) from female C57BL/6J mouse kidney and/or brain genomic DNA. The clone and detailed information can be obtained from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>

NEIGHBORING SEQUENCE INFORMATION:

This sequence is the entire insert of the clone.

FEATURES

source	Location/Qualifiers
	1..200774
	/organism="Mus musculus"
	/mol_type="genomic DNA"
	/db_xref="taxon:10090"
	/chromosome="16"
	/map="16"
	/clone="RP23-282P1"
	/clone_lib="RPCI-23"
repeat_region	52..251
	/rpt_family="B2"
repeat_region	754..1225
	/rpt_family="L1"
repeat_region	1230..1461
	/rpt_family="L1"
repeat_region	1835..2109
	/rpt_family="B4"
repeat_region	2159..2276
	/rpt_family="MaLR"
repeat_region	2301..2473
	/rpt_family="B2"
repeat_region	2665..2741
	/rpt_family="Alu"
repeat_region	2742..2878
	/rpt_family="Alu"
repeat_region	2879..2966
	/rpt_family="Alu"
repeat_region	4183..4325
	/rpt_family="ERVK"
repeat_region	4455..4567
	/rpt_family="ERVK"
repeat_region	4732..4930
	/rpt_family="MaLR"
repeat_region	4954..5144
	/rpt_family="L1"
repeat_region	5132..6177
	/rpt_family="L1"
repeat_region	6178..6398
	/rpt_family="MaLR"
repeat_region	6732..6924
	/rpt_family="B4"
repeat_region	6930..7001
	/rpt_family="B4"
repeat_region	7019..7097
	/rpt_family="ID"
repeat_region	7169..7632
	/rpt_family="L1"
repeat_region	7792..8195
	/rpt_family="MaLR"
repeat_region	8458..8725
	/rpt_family="B4"
repeat_region	8944..8998
	/rpt_family="B2"
repeat_region	8999..9451
	/rpt_family="ERV1"
repeat_region	9452..9592
	/rpt_family="B2"
repeat_region	9932..10075
	/rpt_family="B4"
repeat_region	10765..10875
	/rpt_family="B4"
repeat_region	11883..12045
	/rpt_family="B4"
repeat_region	12128..12498
	/rpt_family="RMER17C"
repeat_region	12546..12596
	/rpt_family="B4"
repeat_region	16707..16777
	/rpt_family="Alu"
repeat_region	17062..17381
	/rpt_family="B4"
repeat_region	19405..19807
	/rpt_family="L1"
repeat_region	19806..20003
	/rpt_family="L1"
repeat_region	21347..21727
	/rpt_family="MaLR"
repeat_region	21887..22098
	/rpt_family="MaLR"
repeat_region	22803..22853
	/rpt_family="B4"
repeat_region	22971..23430
	/rpt_family="ERVK"
repeat_region	23426..23593
	/rpt_family="B4"
repeat_region	23816..23985
	/rpt_family="B4"
repeat_region	23992..24131
	/rpt_family="ERV1"
repeat_region	25013..25148
	/rpt_family="Alu"
repeat_region	25661..25816
	/rpt_family="B4"
repeat_region	26032..26117
	/rpt_family="B4"
repeat_region	26797..26938
	/rpt_family="Alu"
repeat_region	27033..27194
	/rpt_family="L1"
repeat_region	27215..27721
	/rpt_family="L1"
repeat_region	28012..28220
	/rpt_family="ERV1"
repeat_region	28221..28630
	/rpt_family="MaLR"
repeat_region	28631..28777

```

/rpt_family="ERV1"
29116..29394
/rpt_family="L1"
29651..29750
/rpt_family="Alu"
29769..29865
/rpt_family="L1"
30023..30248
/rpt_family="L1"
30249..30661
/rpt_family="ERVK"
30662..30911
/rpt_family="L1"
31018..31183
/rpt_family="B4"
33853..33941
/rpt_family="B4"

Query Match      85.6%; Score 21.4; DB 9; Length 200774;
Best Local Similarity 95.7%; Pred. No. 15;
Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY  2 AAAAACTTGGAGGGAGACTC 24
|||||
Db  84737 AAAAACTTGGAGGGAGACTC 84715

RESULT 4
BC048869
LOCUS      844 bp mRNA linear ROD 08-MAR-2005
DEFINITION Mus musculus kallikrein 9, mRNA (CDNA clone MGC:59025
IMAGE:4164261), complete cds.
ACCESSION BC048869.1 GI:29179502
VERSION    MGC.
KEYWORDS   Mus musculus (house mouse)
SOURCE     Mus musculus
ORGANISM   Mus musculus (house mouse)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.
1 (bases 1 to 844)
Strausberg RL, Feingold EA, Grouse LH, Derge JG, Klausner RD,
Collins FS, Wagner CM, Schuler GD, Altshul SF, Zeeberg
B, Buetow KH, Schaefer CF, Bhat NK, Hopkins RF, Jordan H, Moore T,
Max SI, Wang J, Haieh F, Diatchenko L, Marusina K, Farmer AA, Rubin
GM, Hong L, Stapleton M, Soares MB, Bonaldo MF, Casavant TL,
Schetz TE, Brownstein MJ, Usdin TB, Toshiyuki S, Carninci P,
Prange C, Raha SS, Loquellano NA, Peters GJ, Abramson RD, Mullahy
SJ, Bosak SA, McEwan PJ, McKernan KJ, Malek JA, Gunaratne PH,
Richards S, Worley KC, Hale S, Garcia AM, Gay LJ, Hulyk SW,
Villalon DK, Muzny DM, Sodergren EJ, Lu X, Gibbs RA, Fahey J,
Helton E, Kettman M, Madan A, Rodrigues S, Sanchez A, Whiting M,
Madan A, Young AC, Shevchenko Y, Bouffard GG, Blakesley RW,
Touchman JW, Green ED, Dickson MC, Rodriguez AC, Grimwood J,
Schmutz J, Myers RM, Butterfield YS, Krzywinski MI, Skalska U,
Smalius DE, Schnerch A, Schein JE, Jones SJ and Marra MA.
Mammalian Gene Collection Program Team
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 844)
Director MGC Project.
Direct Submission
Submitted (14-MAR-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: http://www.nisc.nih.gov/
Contact: nisc_mgc@nigri.nih.gov
Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,
Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R.,
Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C.,
McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W.,
Tsurgoun,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,
Young,A., Zhang,L.-H. and Green,E.D.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 108 Row: C Column: 10
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 21426850.

FEATURES
Location/Qualifiers
1..844
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="MGC:59025 IMAGE:4164261"
/tissue_type="Salivary gland, 10 week old female mouse"
/clone_lib="NCI CGAP_SG2"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
1..844
/gene="Klk9"
/note="synonyms: Klklb9, mGk-9"
/db_xref="GeneID:13648"
/db_xref="MGI:95293"
1..786
/gene="Klk9"
/codon_start=1
/product="kallikrein 9"
/protein_id="AAH4869.1"
/db_xref="GI:29179503"
/db_xref="GeneID:13648"
/db_xref="MGI:95293"
/translation="MRFLILFLALSLGIDAAPVHRSIVGGFKCKEKNSQPHWVAVYR
YNEVLCGVLLDANWVLTAAHCYVEENKSLGKNNLYEEPSAQHRLVSKSLFHPGYN
RSLRHHRRHFEDYSNDMLRLSKPADIDIVVXPALPTEEPKLSCTCLASGGST
TPFKFNARDLQCVNKLPLNEDCGKAHIEKVTVDVNLCAETDGGDKTCGDSGGPLI
CDGVLOGITSGFTPCGPKPGVYTKLIKFTSWIKDTMAKNL"

ORIGIN
Query Match      83.2%; Score 20.8; DB 9; Length 844;
Best Local Similarity 91.7%; Pred. No. 62;
Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY  2 AAAAACTTGGAGGGAGACTCA 25
|||||
Db  616 AAGACACTTGCAGGGAGACTCA 639
|||||

RESULT 5
BC024624
LOCUS      859 bp mRNA linear ROD 29-JUN-2004
DEFINITION Mus musculus kallikrein 9, mRNA (CDNA clone MGC:25400
IMAGE:4914236), complete cds.
ACCESSION BC024624
VERSION    MGC.
KEYWORDS   Mus musculus (house mouse)
SOURCE     Mus musculus
ORGANISM   Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.

```


KEYWORDS	MGC.
SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridea; Muridae; Murinae; Mus.
AUTHORS	1 (bases 1 to 878) Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,P., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Schaefer,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carrinci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullaly,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shvchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalls,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
PUBMED	12477932
REFERENCE	2 (bases 1 to 878)
AUTHORS	Strausberg,R.
TITLE	Direct Submission
JOURNAL	Submitted (02-APR-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT	Contact: MGC help desk Email: cgabs@mail.nih.gov Tissue Procurement: Jeffrey E. Green, M.D. CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305 Web site: http://www.shgc.stanford.edu Contact: (Dickson, Mark) mcd@paxil.stanford.edu Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.
FEATURES	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Series: IRAC Plate: 43 Row: d Column: 13 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 8659567. Location/Qualifiers 1..878 /organism="Mus musculus" /mol_type="mRNA" /strain="FVB/N" /db_xref="taxon:10090" /clone="MGC:31128 IMAGE:4165617" /tissue_type="Salivary gland, 10 week old female mouse" /clone_lib="NCI CGAP_SG2" /lab_host="DH10B" /notes="Vector: pCMV-SPORT6"
source	1..878 /gene="Klk1" /notes="synonyms: Klk6, mGK-1, TK, mK1" /db_xref="GeneID:16623" /db_xref="MGI:892019" 20..805 /gene="Klk1" /codon_start=1
gene	AC165233 Mus musculus chromosome 8 clone RP23-30702 map 8, *** SEQUENCING IN PROGRESS ***, 3 unordered pieces. AC165233 AC165233.4 GI:71143318 HTG; HTGS PHASE1; HTGS FULLTOP; HTGS_ACTIVEFIN. Mus musculus (house mouse)
CDS	

Query Match	83.2%;	Score	20.8;	DB	9;	Length	878;
Best Local Similarity	91.7%;	Pred. No.	62;				
Matches	22;	Conservative	0;	Mismatches	2;	Indels	0;
Gaps	0;						

QY	2	AAAAACACTTGGAGGGAGACTCA	25
Db	635	AAAGACACTTGTAAAGGAGACTCA	658

RESULT 8							
AY493302/c							
LOCUS	412 bp DNA linear VRT 30-NOV-2004						
DEFINITION	Anas platyrhynchos microsatellite CAUD057 sequence.						
ACCESSION	AY493302						
VERSION	AY493302.1 GI:40794613						
KEYWORDS	.						
SOURCE	Anas platyrhynchos						
ORGANISM	Anas platyrhynchos						
REFERENCE	Archosauria; Aves; Neognathae; Craniata; Vertebrata; Euteleostomi; Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Craniata; Vertebrata; Euteleostomi; 1 (bases 1 to 412)						
AUTHORS	Huang,Y.H., Tu,J.F., Cheng,X.B., Tang,B., Hu,X.X., Liu,Z.L., Feng,J.D., Lou,Y.K., Lin,L., Xu,K., Zhao,Y.D. and Li,N.						
TITLE	Isolation and characterization of 102 novel microsatellite DNA markers from the duck (Anas platyrhynchos) genome						
JOURNAL	Unpublished						
REFERENCE	2 (bases 1 to 412)						
AUTHORS	Huang,Y.H., Tu,J.F., Cheng,X.B., Tang,B., Hu,X.X., Liu,Z.L., Feng,J.D., Lou,Y.K., Lin,L., Xu,K., Zhao,Y.D. and Li,N.						
TITLE	Direct Submission						
JOURNAL	Submitted (01-DEC-2003) State Key Laboratory for Agrobiotechnology, China Agricultural University, Yuanmingyuan West Road 2, Beijing 100094, China						
FEATURES	Location/Qualifiers 1..412 /organism="Anas platyrhynchos" /mol_type="genomic DNA" /db_xref="taxon:8839"						
repeat_region	1..412 /note="microsatellite CAUD057" /rpt_type=tandem						
ORIGIN							
Query Match	80.8%;	Score	20.2;	DB	5;	Length	412;
Best Local Similarity	88.0%;	Pred. No.	1.3e+02;				
Matches	22;	Conservative	0;	Mismatches	3;	Indels	0;
Gaps	0;						

QY	1	AAAAACACTTGGAGGGAGACTCA	25
Db	304	AAAAACAGATGAAGGGAGACTCA	280

RESULT 9	
AC165233/c	
LOCUS	158857 bp DNA linear HTG 26-JUL-2005
DEFINITION	Mus musculus chromosome 8 clone RP23-30702 map 8, *** SEQUENCING IN PROGRESS ***, 3 unordered pieces.
ACCESSION	AC165233
VERSION	AC165233.4 GI:71143318
KEYWORDS	HTG; HTGS PHASE1; HTGS FULLTOP; HTGS_ACTIVEFIN.
SOURCE	Mus musculus (house mouse)

ORGANISM	Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Murioidea; Muridae; Murinae; Mus.
AUTHORS	1 (bases 1 to 158857)
TITLE	Birren, B., Nusbaum, C. and Lander, E.
REFERENCE	Mus musculus chromosome 8, clone RP23-30702
AUTHORS	Unpublished
REFERENCE	2 (bases 1 to 158857)
AUTHORS	Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N., Anderson, M., Anderson, S., Arachchi, H. M., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgalter, B., Camarata, J., Chang, J., Choepel, Y., Collamore, A., Cook, A., Cooke, P., Corum, B., DeArrellano, K., Diaz, J. S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kanat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Liu, X., Lui, A., Mabbitt, R., MacLean, C., MacDonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nguyen, T., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V. S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
TITLE	Direct Submission
JOURNAL	Submitted (07-JUL-2005) Broad Institute of MIT and Harvard, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE	3 : (bases 1 to 158857)
AUTHORS	Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N., Anderson, M., Anderson, S., Arachchi, H. M., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgalter, B., Camarata, J., Chang, J., Choepel, Y., Collamore, A., Cook, A., Cooke, P., Corum, B., DeArrellano, K., Diaz, J. S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kanat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Liu, X., Lui, A., Mabbitt, R., MacLean, C., MacDonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nguyen, T., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V. S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
TITLE	Direct Submission
JOURNAL	Submitted (26-JUL-2005) Broad Institute of MIT and Harvard, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE	On Jul 26, 2005 this sequence version replaced gi:71061547.
AUTHORS	All repeats were identified using RepeatMasker:
COMMENT	Smit, A.F.A. & Green, P. (1996-1997) http://ftp.genome.washington.edu/RM/RepeatMasker.html
REFERENCE	----- Genome Center
AUTHORS	Center: Broad Institute of MIT and Harvard
TITLE	Center code: WBTR
JOURNAL	Web site: http://www-seq.wi.mit.edu
REFERENCE	Contact: sequence_submissions@broad.mit.edu
AUTHORS	----- Project Information
TITLE	Center project name: L33365
JOURNAL	Center clone name: 307_O_2
REFERENCE	-----
AUTHORS	* NOTE: This is a 'working draft' sequence. It currently consists of 3 contigs. The true order of the pieces

FEATURES	source
1..158857	
/organism="Mus musculus"	
/mol_type="genomic DNA"	
/db_xref="taxon:10090"	
/chromosome="8"	
/map="8"	
/clone="RP23-30702"	
/clone_lib="RPCI-23 Female Mouse BAC"	
44785..44884	
/estimated_length=unknown	
78012..78111	
/estimated_length=unknown	
ORIGIN	
Query Match	80.8%; Score 20.2; DB 14; Length 158857;
Best Local Similarity	88.0%; Pred. No. 59;
Matches	22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY	1 AAAAAACACTTGGAGGGAGACTCA 25
Db	80078 AAAAAACACTTGGAGGGAGACTCA 80054
RESULT 10	
AC142350/c	
LOCUS	AC142350 168803 bp DNA linear PRI 18-SEP-2003
DEFINITION	Pan troglodytes BAC clone RP43-12B4 from Y, complete sequence.
ACCESSION	AC142350
VERSION	AC142350.1 GI:29294469
KEYWORDS	HTG.
SOURCE	Pan troglodytes (chimpanzee)
ORGANISM	Pan troglodytes
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Pan.
AUTHORS	1 (bases 1 to 168803)
TITLE	Isak, A. and Haakenson, W.
JOURNAL	The sequence of Pan troglodytes BAC clone RP43-12B4
REFERENCE	Unpublished (2001)
AUTHORS	2 (bases 1 to 168803)
TITLE	Suleston, J. E. and Wilson, R.
JOURNAL	Sequencing of Pan troglodytes
REFERENCE	Unpublished (2001)
AUTHORS	3 (bases 1 to 168803)
TITLE	Wilson, R. K.
JOURNAL	Direct Submission
REFERENCE	Submitted (27-MAR-2003) Genetics, Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
AUTHORS	4 (bases 1 to 168803)
TITLE	Wilson, R.
JOURNAL	Direct Submission
REFERENCE	Submitted (18-SEP-2003) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
AUTHORS	----- Genome Center
TITLE	Center: Washington University Genome Sequencing Center
JOURNAL	Center code: WUGSC
REFERENCE	Web site: http://genome.wustl.edu
AUTHORS	Contact: submissions@wustl.edu
TITLE	----- Summary Statistics
JOURNAL	Center project name: C_PT012B04
REFERENCE	-----

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

The position of this clone was established as part of a collaboration between the Chimpanzee Chromosome Y Mapping Project (Jennifer F. Hughes, Tatyana Pyntikova, Helen Skaletsky, Steve Rozen, and David C. Page at the Whitehead Institute for Biomedical Research, Cambridge MA) and the Washington University Genome Sequencing Center, St. Louis MO.

SOURCE INFORMATION:

The RPCT-43 BAC library has been constructed by Chung-Li Shu. DNA was isolated from white blood cells obtained from a male chimpanzee (Pan troglodytes, 'Clint', Yerkes #C0471; birthdate: 6-6-80). The clone and detailed information can be obtained from ResGen (<http://www.resgen.com>) or Pieter de Jong and co-workers at <http://www.bacpac.chori.org>.

NEIGHBORING SEQUENCE INFORMATION:

This sequence is the entire insert of the clone.

FEATURES
source

Location/Qualifiers
1..168803
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/chromosome="Y"
/map="Y"
/clone="RP43-12B4"
/clone_lib="RPCT-43"

ORIGIN

Query Match 80.8%; Score 20.2; DB 8; Length 168803;
Best Local Similarity 88.0%; Pred. No. 58;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAACACTTGGAGGGAGACTCA 25

Db 99854 AAAAAACATGGAGAGATACTCA 99830

RESULT 11
AC147666/c 189158 bp DNA linear PRI 30-JUN-2004
LOCUS Pan troglodytes BAC clone CH251-19607 from Y, complete sequence.
DEFINITION
AC147666
VERSION
AC147666.4 GI:49458043
KEYWORDS
HTG.

SOURCE
ORGANISM
Pan troglodytes (chimpanzee)
Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Pan.

REFERENCE
1 (bases 1 to 189158)
Belter, E., Haglund, K., Elliott, G. and Meyer, R.
AUTHORS The sequence of Pan troglodytes BAC clone CH251-19607

TITLE Unpublished (2001)
JOURNAL
REFERENCE 2 (bases 1 to 189158)
AUTHORS Wilson, R.K.

TITLE Direct Submission
JOURNAL
Submitted (21-DEC-2003) Genetics, Genome Sequencing Center, 4444
Forest Park Parkway, St. Louis, MO 63108, USA

REFERENCE
AUTHORS
TITLE
JOURNAL

REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT

3 (bases 1 to 189158)

Wilson, R.K.
Direct Submission
Submitted (24-APR-2004) Genetics, Genome Sequencing Center, 4444
Forest Park Parkway, St. Louis, MO 63108, USA

4 (bases 1 to 189158)

Wilson, R.K.
Direct Submission
Submitted (30-JUN-2004) Washington University School of Medicine,
Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO
63108, USA

On Jun 30, 2004 this sequence version replaced gi:46395453.

----- Genome Center

Center: Washington University Genome Sequencing Center

Center code: WUGSC

Web site: <http://genome.wustl.edu>

Contact: submissions@watson.wustl.edu

----- Summary Statistics

----- Center project name: C_AB0196007

NOTICE:

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

The position of this clone was established as part of a collaboration between the Chimpanzee Chromosome Y Mapping Project (Jennifer F. Hughes, Tatyana Pyntikova, Helen Skaletsky, Steve Rozen, and David C. Page at the Whitehead Institute for Biomedical Research, Cambridge MA) and the Washington University Genome Sequencing Center, St. Louis MO.

SOURCE INFORMATION:

The CHORI-251 Chimpanzee BAC library has been constructed at the Children's Hospital Oakland Research Institute, BACPAC Resources, by Dr. Baoli Zhu. DNA was isolated from white blood cells obtained from a male chimpanzee (Pan troglodytes, 'Clint', Yerkes #C0471; birthdate: 6-6-80). The clone and detailed information can be obtained from Pieter de Jong and co-workers at <http://www.bacpac.chori.org>.

NEIGHBORING SEQUENCE INFORMATION:

This sequence is the entire insert of the clone. This clone is overlapped by AC147127 and AC146455.

FEATURES
source

Location/Qualifiers
1..189158
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/chromosome="Y"
/map="Y"
/clone="CH251-19607"
/clone_lib="CHORI251"
29362..29508

misc_feature

/note="Sequence derived from PCR product of project DNA."

unsure

/note="Sequence derived from one plasmid subclone."

misc_feature

/note="Sequence derived from PCR product of project DNA."

ORIGIN

Query Match 80.8%; Score 20.2; DB 8; Length 189158;

Best Local Similarity 88.0%; Pred. No. 58;

Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAACACTTGGAGGGAGACTCA 25

```

|||||
170949 AAAAAACATCGAAGAGACTCA 170925

RESULT 12
LOCUS AC145502 196620 bp DNA linear HTG 08-AUG-2003
DEFINITION Canis familiaris clone RP81-204N10, WORKING DRAFT SEQUENCE.
ACCESSION AC145502
VERSION AC145502.1 GI:32964888
KEYWORDS HTG; HTGS PHASE2; HTGS DRAFT.
SOURCE Canis familiaris (dog)
ORGANISM Canis familiaris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;
Canis.

REFERENCE
1 (bases 1 to 196620)
Antoniellis, A., Ayelle, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S., Cariaga, K.,
Chu, G., Coleman, B., Coleman, H., Engle, J., Granite, S., Guan, X.,
Gupta, J., Haghighi, P., Han, J., Hansen, N., Ho, S.-L., Hu, P.,
Hurle, B., Idol, J.R., Karlins, E., Kwong, P., Lalic, P., Lee-Lin, S.-O.,
Legaspi, R., Maduro, Q.L., Maduro, V.B., Margulies, E.H., Mastello, C.,
Maskeri, B., McDowell, J., Paquirigan, C., Pearson, R., Portnoy, M.E.,
Prasad, A., Reddix-Dugue, N., Schandler, K., Schueler, M.G., Shah, K.,
Sison, C., Stantripop, S., Thomas, J.W., Thomas, P.J., Tsipouri, V.,
Vogt, J.L., Wetherby, K.D., Young, A. and Green, E.D.

TITLE
NISC Comparative Sequencing Initiative
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 196620)
Green, E.D.
AUTHORS
Direct Submission
JOURNAL
Submitted (18-JUL-2003) NIH Intramural Sequencing Center, 8717
Grovemont Circle, Gaithersburg, MD 20877, USA
3 (bases 1 to 196620)
Green, E.D.
AUTHORS
Direct Submission
JOURNAL
Submitted (08-AUG-2003) NIH Intramural Sequencing Center, 8717
Grovemont Circle, Gaithersburg, MD 20877, USA
COMMENT
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc_zoo@hgri.nih.gov
----- Project Information
Center project name: etm
Center clone name: 204N10

The sequence data in this record represents an 'enhanced'
version of a Phase 2 submission. Specifically, the indicated
order and orientation of each sequence contig has been
established using one or more of the following: read-pair
data from individual subclones, overlaps with neighboring
clones, alignment with available reference sequence (e.g.,
human), and/or confirmation by PCR testing. In addition,
the sequence assembly is based on at least 8X average
coverage in Q20 bases and has been reviewed to rule out
gross misassemblies, the low-quality ends of sequence
contigs have been trimmed away, and each base is associated
with a Phrap-derived quality score.
----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 196374 bases at least Q40
Consensus quality: 196562 bases at least Q30
Consensus quality: 196601 bases at least Q20
Insert size: 185000; agarose-fp
Insert size: 196620; sum-of-contigs
Quality coverage: 10.83x in Q20 bases; agarose-fp
Quality coverage: 10.19x in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently

```

```

* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 196620: contig of 196620 bp in length.

FEATURES
Location/Qualifiers
1..196620
/organism="Canis familiaris"
/mol_type="genomic DNA"
/db_xref="taxon:9615"
/clone="RP81-204N10"
/clone_lib="RP81"
1..196620
/notes="assembly_fragment"
clone_end:17
vector_side:left
clone_end:Sp6
vector_side:right

ORIGIN
Query Match 80.8%; Score 20.2; DB 14; Length 196620;
Best Local Similarity 88.0%; Pred. No. 57;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAACATCGAAGGAGACTCA 25
Db 185805 ACAACACACTTGCAGGAGACTCA 185829

RESULT 13
AC102363 210317 bp DNA linear ROD 22-JAN-2005
LOCUS Mus musculus chromosome 1, clone RP23-76B18, complete sequence.
DEFINITION AC102363
ACCESSION AC102363.19 GI:58036679
VERSION HTG.
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.

REFERENCE
1 (bases 1 to 210317)
Birren, B., Nusbaum, C. and Lander, E.
Mus musculus chromosome 1, clone RP23-76B18
Unpublished
REFERENCE
2 (bases 1 to 210317)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N.,
Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Boukhgalter, B.,
Brown, A., Camarata, J., Campopiano, A., Chang, J., Chararo, B.,
Choepe, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A.,
Cooke, P., DeArelano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S.,
Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S.,
Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
Hagos, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,
Jones, C., Kamat, A., Karatas, A., Kells, C., Lakocque, K.,
Lamazzaro, R., Landers, T., Lehoczy, J., Levine, R., Liu, G.,
MacLean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C.,
McCarthy, M., McEwan, P., McKernan, K., McPheeters, R., Meidrim, J.,
Meneus, L., Mihova, T., Menga, V., Murphy, T., Naylor, J., Nguyen, C.,
Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D.,
Olivier, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,
Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupbach, R.,
Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
Topham, K., Travers, M., Travis, N., Trigglio, J., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
Direct Submission
TITLE

```

JOURNAL

REFERENCE

AUTHORS

Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 3 (bases 1 to 210317)
 Birren,B., Nubaum,C., Lander,E., Abouelleil,A., Allen,N., Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B., DeArellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N., Hagoopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Katat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbitt,R., MacLean,C., Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nguyen,T., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE

JOURNAL

REFERENCE

AUTHORS

Submitted (19-OCT-2004) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 4 (bases 1 to 210317)
 Birren,B., Nubaum,C., Lander,E., Abouelleil,A., Allen,N., Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B., DeArellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N., Hagoopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Katat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbitt,R., MacLean,C., Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nguyen,T., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE

JOURNAL

COMMENT

Submitted (22-JAN-2005) Broad Institute of MIT and Harvard, 320 Charles Street, Cambridge, MA 02141, USA
 On Jan 22, 2005 this sequence version replaced gi:54291951.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Broad Institute of MIT and Harvard

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@broad.mit.edu

----- Project Information

Center project name: L18576

Center clone name: 76_B_18

FEATURES

source

Location/Qualifiers

1..210317
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /chromosome="1"
 /map="1"
 /clone="RP23-76B18"

/clone_lib="RPCI-23 Female Mouse BAC"
 complement(55..1188)
 /rpt_family="Lx4"
 1307..1368
 /rpt_family="GA)n"
 1465..1645
 /rpt_family="Lx8"
 1855..2172
 /rpt_family="Lx8"
 complement(2450..2823)
 /rpt_family="MTA"
 3335..4262
 /rpt_family="Lx2B"
 4263..4287
 /rpt_family="AT_rich"
 4385..4460
 /rpt_family="TA)n"
 complement(5005..5347)
 /rpt_family="ORR1B1"
 5348..5376
 /rpt_family="(T)n"
 complement(5401..5717)
 /rpt_family="L1_MM"
 5769..5819
 /rpt_family="L1M2"
 6130..6170
 /rpt_family="AT_rich"
 6648..6826
 /rpt_family="B3"
 6901..7087
 /rpt_family="Lx"
 7098..7141
 /rpt_family="(TAA)n"
 7142..7166
 /rpt_family="(TTA)n"
 7168..7987
 /rpt_family="Lx"
 10295..10670
 /rpt_family="ORR1B2"
 11203..11253
 /rpt_family="(TAAA)n"
 11297..11394
 /rpt_family="AT_rich"
 complement(12397..12888)
 /rpt_family="L1_MM"
 13083..13745
 /rpt_family="RMER6A"
 14619..15735
 /rpt_family="L1VL2"
 15736..15821
 /rpt_family="(TA)n"
 15822..19354
 /rpt_family="L1VL2"
 19342..19713
 /rpt_family="L1_MM"
 19708..21263
 /rpt_family="Lx6"
 21256..22097
 /rpt_family="Lx"
 complement(22304..22464)
 /rpt_family="MT2A"
 23133..23184
 /rpt_family="(TA)n"
 23826..23854
 /rpt_family="(A)n"
 24847..24980
 /rpt_family="B4"
 25004..25038
 /rpt_family="(TCTG)n"
 25040..25064
 /rpt_family="(TA)n"
 25370..25395
 /rpt_family="(TTTG)n"

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

HTG 23-JAN-2005

HTG 23-JAN-2005

DEFINITION	Bos taurus clone CH240-4L7, WORKING DRAFT SEQUENCE, 20 unordered pieces.
ACCESSION	AC152260
VERSION	AC152260.2 GI:58038041
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE	Bos taurus (cow)
ORGANISM	Bos taurus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.
AUTHORS	Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-osman, F.R., Allen, C., Albrooks, S.L., Anarstange, H.C., Are, J.R., Ayele, M., Banks, T., Barbara, J., Benton, J., Binage, K., Blankenburg, K., Bonnin, D., Bouck, J., Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P., Buhay, C., Burch, P., Burrell, C., Burrell, K.L., Byrd, N.C., Chen, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C., Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J., Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hogue, M., Holloway, C., Hollins, B., Homs, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Loulseghe, H., Lozado, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E., Massey, E., Mawhney, E., McLeod, M.P., Meador, M., Mei, G., Metzker, M., Moser, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S., Nguyen, N., Nickerson, B., Nwokenkwo, S., Ogih, M., Okwuonu, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savery, G., Scherer, S., Scott, G., Shen, H., Shoostari, N., Sisson, I., Sodergren, E., Sonaik, R., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Usmani, K., Vasquez, L., Vera, V., Verduzco, D., Villalón, D., Vinson, R., Waldron, L.M., Wang, Q., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S., Williams, G., Williamson, A., Wleczka, R., Woodson, S., Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D., Weinstock, G. and Gibbs, R.
TITLE	Direct Submission
REFERENCE	Unpublished
AUTHORS	2 (bases 1 to 215892)
TITLE	Worley, K.C.
JOURNAL	Direct Submission
AUTHORS	Submitted (01-NOV-2004) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
JOURNAL	3 (bases 1 to 215892)
REFERENCE	Worley, K.C.
AUTHORS	Direct Submission
JOURNAL	Submitted (23-JAN-2005) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
COMMENT	On Jan 23, 2005 this sequence version replaced gi:55058582. The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence

may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: FAFR
Center clone name: CH240-4L7
----- Summary Statistics
Assembly program: Atlas 3.0;
Consensus quality: 209512 bases at least Q40
Consensus quality: 211147 bases at least Q30
Consensus quality: 212594 bases at least Q20
Estimated insert size: 211936; sum-of-contigs estimation
Quality coverage: 5x in Q20 bases; sum-of-contigs estimation

*** NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 20 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
* as soon as it is available and the accession number will
* be preserved.

1	1232:	contig of 1232 bp in length
1233	1282:	gap of 50 bp
1283	20143:	contig of 18661 bp in length
20144	20193:	gap of 50 bp
20194	28382:	contig of 8189 bp in length
28383	28432:	gap of 50 bp
28433	50691:	contig of 22259 bp in length
50692	50791:	gap of unknown length
50792	85570:	contig of 34779 bp in length
85571	85620:	gap of 50 bp
85621	90790:	contig of 5170 bp in length
90791	90929:	gap of 139 bp
90930	95104:	contig of 4175 bp in length
95105	95154:	gap of 50 bp
95155	102719:	contig of 7565 bp in length
102720	102769:	gap of 50 bp
134605	134704:	contig of 31835 bp in length
134705	137956:	contig of 3252 bp in length
137957	138006:	gap of 50 bp
138007	153576:	contig of 15570 bp in length
153577	153626:	gap of 50 bp
153627	162660:	contig of 9034 bp in length
162661	163236:	gap of 576 bp
163237	198080:	contig of 34844 bp in length
198081	198405:	gap of 325 bp
198406	200117:	contig of 1712 bp in length
200118	200167:	gap of 50 bp
200168	207948:	contig of 7781 bp in length
207949	207998:	gap of 50 bp
207999	209885:	contig of 1887 bp in length
209886	209985:	gap of unknown length
209986	211031:	contig of 1046 bp in length
211032	211131:	gap of unknown length
211132	212180:	contig of 1049 bp in length
212181	212280:	gap of unknown length
212281	213281:	contig of 1048 bp in length
213282	213428:	gap of unknown length
213429	215892:	contig of 2464 bp in length.
213429		Location/Qualifiers
	1.	215892
		/organism="Bos taurus"

FEATURES
source


```

/mol_type="genomic DNA"
/db_xref="taxon:9913"
/c1one="CH240-4L7"
1233..1282
/estimated_length=50
20144..20193
/estimated_length=50
28383..28432
/estimated_length=50
50892..50791
/estimated_length=unknown
85571..85620
/estimated_length=50
90791..90929
/estimated_length=139
95105..95154
/estimated_length=50
102720..102769
/estimated_length=50
134605..134704
/estimated_length=unknown
137957..138006
/estimated_length=50
153577..153626
/estimated_length=50
162661..163236
/estimated_length=576
198081..198405
/estimated_length=325
200118..200167
/estimated_length=50
207949..207998
/estimated_length=50
209886..209985
/estimated_length=unknown
211032..211131
/estimated_length=unknown
212181..212280
/estimated_length=unknown
213329..213428
/estimated_length=unknown

```

ORIGIN

```

Query Match      80.8%; Score 20.2; DB 14; Length 215892;
Best Local Similarity 88.0%; Pred. No. 57;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

Qy 1 AAAAAACACTTGGAGGAGACTCA 25
    ||||| ||||| ||||| ||||| |||||
Db 108402 AAAAAACCTAGAGGAGACTCA 108378

```

```

RESULT 16
AC116092 Rattus norvegicus clone CH230-134D13, *** SEQUENCING IN PROGRESS
LOCUS AC116092 230954 bp DNA linear HTG 08-OCT-2002
DEFINITION ***; 11 unordered pieces.
AC116092
VERSION AC116092.3 GI:23269875
KEYWORDS HTGS PHASE1; HTGS DRAFT; HTGS_ENRICHED.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus

```

```

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Rattus.
1 (bases 1 to 230954)
Murny D.Marie., Metzker M.Lee., Abramson S., Adams C., Alder J.,
Allen C., Allen H., Alsbrooks S., Amin A., Anguiano D.,
Anyalebechi V., Aoyagi A., Ayodeji M., Baca E., Baden H.,
Baldwin D., Bandaranaike D., Barber M., Barnstead M., Benahmed F.,
Biswal K., Blair J., Blankenburg K., Blyth P., Brown M.,
Bryant N., Buhay C., Burch P., Burrell K., Calderon E.,
Cardenas V., Carter K., Cavazos I., Ceasar H., Center A.,

```

Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escoto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Georegeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogue, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensu, L., Loubege, H., Lozano, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhiney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Parks, K., Nwaokelimeh, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Polindexter, A., Popovic, D., Primus, E., Pu, L., L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajs, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villabana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wlarczyk, R., Woodden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, X., Zhao, S., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G., and Gibbs, R.A.

TITLE

Unpublished

2 (bases 1 to 230954)

Worley, K.C.

Direct Submission

TITLE

JOURNAL

Submitted (24-MAR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 230954)

Rat Genome Sequencing Consortium.

Direct Submission

TITLE

JOURNAL

Submitted (08-OCT-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

COMMENT

On Sep 23, 2002 this sequence version replaced gi:21745909. The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: CH230-134D13
 ----- Summary Statistics
 Assembly program: Phrap; version 0.990329
 Consensus quality: 174870 bases at least Q40
 Consensus quality: 178840 bases at least Q30
 Consensus quality: 181745 bases at least Q20
 Estimated insert size: 199572; sum-of-contigs estimation
 Quality coverage: 3x in Q20 bases; sum-of-contigs estimation

*** NOTE: Estimated insert size may differ from sequence length
 (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 11 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence.
 * as soon as it is available and the accession number will
 * be preserved.

1 4497: contig of 4497 bp in length
 * 4498 4537: gap of unknown length
 * 4598 25239: contig of 20702 bp in length
 * 25300 25399: gap of unknown length
 * 25400 208371: contig of 182972 bp in length
 * 208372 208471: gap of unknown length
 * 208472 209709: contig of 1238 bp in length
 * 209710 209809: gap of unknown length
 * 209810 210979: contig of 1170 bp in length
 * 210980 211079: gap of unknown length
 * 211080 212892: contig of 1813 bp in length
 * 212893 212992: gap of unknown length
 * 212993 218147: contig of 5155 bp in length
 * 218148 218247: gap of unknown length
 * 218248 219391: contig of 1144 bp in length
 * 219392 219491: gap of unknown length
 * 219492 225125: contig of 5634 bp in length
 * 225126 225225: gap of unknown length
 * 225226 227317: contig of 2092 bp in length
 * 227318 227417: gap of unknown length
 * 227418 230954: contig of 3537 bp in length.

FEATURES
 source
 1. .230954
 Location/Qualifiers
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-134D13"

misc_feature
 1. .1265
 /note="wgs contig"
 gap
 4498. .4597
 /estimated_length=unknown
 misc_feature
 21796. .22828
 /note="wgs contig"
 gap
 25300. .25399
 /estimated_length=unknown
 misc_feature
 138428. .139768
 /note="wgs contig"
 gap
 208372. .208471
 /estimated_length=unknown
 gap
 209710. .209809
 /estimated_length=unknown
 gap
 210980. .211079
 /estimated_length=unknown
 gap
 212893. .212992
 /estimated_length=unknown
 gap
 218148. .218247
 /estimated_length=unknown
 gap
 219392. .219491
 /estimated_length=unknown
 gap
 225126. .225225
 /estimated_length=unknown

gap
 227318. .227417
 /estimated_length=unknown

ORIGIN

Query Match 80.8%; Score 20.2; DB 14; Length 230954;
 Best Local Similarity 88.0%; Pred. No. 56;
 Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 AAAAAACACTTGGAGGGAGACTCA 25
 |||||
 Db 110889 AAAAAACACTTGGAGGGAGACTGA 110913

RESULT 17

AC140677 121010 bp DNA linear PRI 27-FEB-2003
 LOCUS Homo sapiens chromosome 5 clone CTC-778L3, complete sequence.
 DEFINITION AC140677
 ACCESSION AC140677
 VERSION AC140677.1 GI:28570298
 KEYWORDS HTG.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Hominidae; Homo.

REFERENCE

1 (bases 1 to 121010)
 AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
 TITLE Direct Submission
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 121010)
 AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
 TITLE Direct Submission
 JOURNAL Submitted (27-FEB-2003) DOE Joint Genome Institute, 2800 Mitchell
 Drive, Walnut Creek, CA 94598, USA
 COMMENT Draft Sequence Produced by DOE Joint Genome Institute
 www.igi.doe.gov
 Finishing Completed at Stanford Human Genome Center
 www.shgc.stanford.edu
 Quality: Phrap Quality >=40 100% of Sequence;
 Estimated Total Number of Errors is 0.

FEATURES

source
 1. .121010
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosomes="5"
 /clone="CTC-778L3"

ORIGIN

Query Match 79.2%; Score 19.8; DB 8; Length 121010;
 Best Local Similarity 91.3%; Pred. No. 95;
 Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 AAAAAACACTTGGAGGGAGACT 23
 |||||
 Db 6013 AAAAAAATTTGGAGGGAGACT 6035

RESULT 18

AC114959 163805 bp DNA linear PRI 22-MAY-2002
 LOCUS Homo sapiens chromosome 5 clone RP11-170L13, complete sequence.
 DEFINITION AC114959
 ACCESSION AC114959
 VERSION AC114959.2 GI:21070666
 KEYWORDS HTG.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Hominidae; Homo.

REFERENCE

1 (bases 1 to 163805)
 AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
 TITLE Direct Submission

JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 163805)
AUTHORS	DOE Joint Genome Institute..
TITLE	Direct Submission
JOURNAL	Submitted (14-MAR-2002) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE	3 (bases 1 to 163805)
AUTHORS	DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE	Direct Submission
JOURNAL	Submitted (22-MAY-2002) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT	On May 22, 2002 this sequence version replaced gi:19424420. Draft Sequence Produced by DOE Joint Genome Institute www.jgi.doe.gov Finishing Completed at Stanford Human Genome Center www.shgc.stanford.edu Quality: Phrap Quality >=40 100% of Sequence; Estimated Total Number of Errors is 0.
FEATURES	Location/Qualifiers 1..163805 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" /chromosome="5" /clone="RP11-170L13"
ORIGIN	Query Match 79.2%; Score 19.8; DB 8; Length 163805; Best Local Similarity 91.3%; Pred No. 92; Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY	1 AAAAAACACTTGGAGGGGAGACT 23 5076 AAAAAAATTGGAAGGGGAGACT 5098
DB	AC156954 179140 bp DNA linear ROD 30-MAY-2005 Mus musculus 6 BAC RP24-343M7 (Roswell Park Cancer Institute (CS7BL/6J Male) Mouse BAC Library) complete sequence. AC156954 AC116805 AC156954.5 GI:66793675 HTG.
ACCESSION	Mus musculus (house mouse)
VERSION	Mus musculus
KEYWORDS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
SOURCE	1 (bases 1 to 179140) Muzny, D., Adams, C., Agbai II, O., Allen, C., Alsbrooks, S., Archer, P., Arradondo, H., Bandaranaike, D., Bangura, L., Beltran, B., Beltran, R., Beraducci, A., Biswal, K., Blythe, P., Bonham, H., Buhay, C., Burch, P., Cadoree, I., Canada, A., Cardenas, V., Carter, K., Cavazos, I., Chacko, J., Chahrour, M., Chavez, D., Chen, A., Chen, G., Chen, R., Cheng, M.-T., Chu, J., Clerc, K., Cockrell, R., Coyle, M., Cree, A., Curry, S., Dai, W., Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Donlin, J., McCauley, S., Dugan-Rocha, S., Dunn, A., Durbin, K., Friunda, D., Egan, A., Escotto, M., Espinosa, V., Eugene, C., Fa, M., Fernandez, S., Fernando, P., Flagg, N., Forbes, L., Foster, P., Fowler, G., Fu, Q., Fuh, E., Garcia, A., Garcia, R., Garner, T., Gasikin, C., Gench, S., Ghose, S., Gill, R., Gonzalez, D., Gonzalez-Garay, M., Guevara, W., Holder, M., Haaland, W., Haeblerlen, K., Hall, B., Hamid, H., Hamilton, K., Harbes, B., Harris, R., Havlak, P., Hawes, A., Hawkins, E., Hayes, S., Hemphill, L., Hernandez, J., Hines, S., Hitchens, M., Hodgson, A., Hogues, M., Hollins, B., Howell, L.T., Hulyk, S., Hume, J., Imo, K., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Kalafus, K., Kelly, S., Keys, T., Khan, Z., King, L., Kovac, C., Kowis, A., Kowis, C., Lara, P., Leal, S., Lee, K., Lee, S., Legall, F.I., Lemon, S., Lewis, L., Li, B., Li, Y., Li, Z., Linell, M., Liu, W., Liu, Y.-S., Liu, Y., Liyanage, D., London, P., Lopez, J., Lorensheuwala, L., Lozano, R., Luk, T., Madu, R.,
REFERENCE	Maheshwari, M., McPherson, J., Mercadao, C., Metzker, M., McLelland, H., Milosavljevic, A., Minja, E., Morgan, M., Morris, S., Munidasa, M., Murray, D., Nazareth, L., Ngo, D., Nguyen, N., Norwig-Eastaugh, E., Nott, A., Nwaokemeleh, O., Obregon, M., Ochi-Okorie, C., Odeh, E., Okwuonu, G., Okwuonu, K., Parker, D., Pasternak, S., Patel, B., Patel, V., Paul, H., Perez, A., Perez, L., Petrosino, J., Pham, T., Primus, E., Pu, L.-L., Puazo, M., Qin, X., Quinn, A., Quiroz, J., Rabata, D., Radcliff, E., Reigh, R., Ren, Y., Reuter, M., Richards, S., Rivers, C., Rodriguez, F., Rojas, A., Ruiz, S.J., Sama, M., Sanders, W., Santibanez, J., Santos, R., Savery, G., Scherer, S., Shen, H., Shen, Y., Sisason, I., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Svatek, A., Taylor, E., Taylor, T., Thomas, N., Thorner, R., Thornton, R., Trejos, Z., Usmani, K., Varco, C., Verduzco, D., Villasana, D., Virk, D., Volkov, A., Waldron, L., Walker, B., Wang, Q., Wang, S., Warren, J., Wei, X., Wheeler, D., Williams, G., Williams, R., Worley, K., Wright, R., Wu, J., Yakub, S., Yan, K., Yuan, Y., Yu, F., Zhang, J., Zhang, L., Zhang, Z., Zhou, J., Weinstock, G. and Gibbs, R.
AUTHORS	Worley, K.C.
TITLE	Direct Submission
JOURNAL	Submitted (01-APR-2005) Human Genome Sequencing Center, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE	4 (bases 1 to 179140)
AUTHORS	Worley, K.C.
TITLE	Direct Submission
JOURNAL	Submitted (30-MAY-2005) Human Genome Sequencing Center, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
COMMENT	On May 30, 2005 this sequence version replaced gi:62122379. Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality does not meet this standard, it will be indicated in the annotation. The repeat regions shown were identified using RepeatMasker by Adrian Smit. Sequence similarities were identified using Powerblast by Jinghui Zhang. Exon/Intron boundaries of identified genes were chosen if there were canonical splice junctions that maintained sequence continuity across the splice junctions. ----- Genome Center Center: Baylor College of Medicine Center code: BCM Web site: http://www.hgsc.bcm.tmc.edu/ Drafting Center Code: WIBR Contact: hgsc-help@bcm.tmc.edu. Location/Qualifiers 1..179140 /organism="Mus musculus" /mol_type="genomic DNA" /db_xref="taxon:10090" /chromosome="6" /clone="RP24-343M7" 1..86096 /note="overlaps bases 133227..219322 of clone AC153628" /function="clone overlap" 9..118 /rpt_family="MTD" 119..199 /rpt_family="(GAAAA)n" 200..260 /rpt_family="MTD"

repeat_region complement(956..1176)
/rpt_family="MTD"
1177..1222
/rpt_family="TGGA)n"
repeat_region complement(1223..1412)
/rpt_family="MTD"
1919..2136
/rpt_family="BAA"
repeat_region 2137..2164
/rpt_family=" (CAAAA)n"
repeat_region 2180..2330
/rpt_family="GA-rich"
repeat_region 2510..2549
/rpt_family=" (GAAAA)n"
repeat_region 2637..2692
/rpt_family=" (CA)n"
repeat_region 2693..2751
/rpt_family=" (GA)n"
repeat_region 2771..2847
/rpt_family="ID"
repeat_region 2848..2876
/rpt_family=" (CAAT)n"
repeat_region 2971..3022
/rpt_family="MTD-int"
repeat_region 3023..3068
/rpt_family=" (TA)n"
repeat_region 3070..3094
/rpt_family=" (GA)n"
repeat_region 3095..3143
/rpt_family="MTD-int"
repeat_region 3145..3696
/rpt_family="Lx7"
repeat_region 3697..3716
/rpt_family=" (TTGG)n"
repeat_region 3717..3867
/rpt_family="Lx7"
repeat_region 3949..4205
/rpt_family="MTD-int"
repeat_region 4213..4349
/rpt_family="B1_Mm"
repeat_region 4351..4436
/rpt_family="A-rich"
repeat_region 4585..5259
/rpt_family="MTD-int"
repeat_region 5277..6562
/rpt_family="L1Md_F"
repeat_region complement(6563..6668)
/rpt_family="L1Md_F"
repeat_region complement(6679..6890)
/rpt_family="ORR1D-int"
repeat_region 6891..9247
/rpt_family="L1_Mm"
repeat_region 9248..9310
/rpt_family="L1_Mus1"
repeat_region 9312..9332
/rpt_family="AT-rich"
repeat_region 9335..9435
/rpt_family="MTD-int"
repeat_region 9450..9771
/rpt_family="MTE"
repeat_region complement(9808..10047)
/rpt_family="Arthuri1"
repeat_region 11734..11838
/rpt_family=" (CAGA)n"
repeat_region complement(11936..12062)
/rpt_family="RSINEL1"
repeat_region complement(12120..12278)
/rpt_family="Mrb"
repeat_region complement(12663..12839)
/rpt_family="B4A"
repeat_region complement(12865..13472)
/rpt_family="Lx3B"
repeat_region complement(14128..14290)

repeat_region /rpt_family="MERSB"
15935..15956
/rpt_family="AT-rich"
repeat_region 17347..17392
/rpt_family=" (TG)n"
repeat_region complement(17607..18001)
/rpt_family="MLTID"
repeat_region 18371..18652
/rpt_family="L2"
repeat_region complement(19155..19548)
/rpt_family="ORR1C1"
repeat_region 20910..21032
/rpt_family="B1_Mus1"
repeat_region complement(21556..22056)
/rpt_family="L1Md_Gf"
repeat_region 22057..22583
/rpt_family="L1Md_T"
repeat_region 22617..22640
/rpt_family=" (A)n"

Query Match 79.2%; Score 19.8; DB 9; Length 179140;
Best Local Similarity 91.3%; Pred. No. 90;
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 AAACACTTGAAGGAGACTCA 25
|||||
Db 74776 AAAGTCITGGAAGGAGACTCA 74754

RESULT 20
AC131565/c
LOCUS AC131565 181283 bp DNA linear PRI 27-AUG-2002
DEFINITION Homo sapiens chromosome 5 clone RP11-1273L8, complete sequence.
ACCESSION AC131565
VERSION AC131565.1 GI:22474846
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 181283)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 181283)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Submitted (24-AUG-2002) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE 3 (bases 1 to 181283)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Submitted (27-AUG-2002) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www-shgc.stanford.edu
Quality: Phrap Quality >=40 99.9% of Sequence;
Estimated Total Number of Errors is 0.1.

FEATURES
source
1..181283
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="5"
/clone="RP11-1273L8"

ORIGIN

Query Match 79.2%; Score 19.8; DB 8; Length 181283;
Best Local Similarity 91.3%; Pred. No. 90;
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAACACTTGAAGGAGACT 23
|||||
Db 177390 AAAAAAATTTGAAGGAGACT 177368

RESULT 21

AF129075/c

LOCUS

DEFINITION

Homo sapiens chromosome 21 clone RP1-100J12, RP1-P79E4 map

q21.3-22.11, complete sequence.

ACCESSION

AF129075

VERSION

AF129075.3 GI:33620784

KEYWORDS

HTG.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominiidae; Homo.

REFERENCE

1 (bases 1 to 190837)

Hattori,M., Fujiyama,A., Taylor,T.D., Watanabe,H., Yada,T.,

Faik,H.-S., Toyoda,A., Ishii,K., Totoki,Y., Choi,D.-K., Soeda,E.,

Ohki,M., Takagi,T., Sakaki,Y., Taudien,S., Blechschmidt,K.,

Polley,A., Menzel,U., Delabar,J., Kumpf,K., Lehmann,R.,

Patterson,D., Reichwald,K., Rump,A., Schillhabel,M.B., Schudy,A.,

Zimmermann,W., Rosenthal,A., Kudoh,J., Kawasaki,K., Asakawa,S.,

Shintani,A., Sasaki,T., Nagamine,K., Mitsuyama,S.,

Antonarakis,S.E., Minoshima,S., Shimizu,N., Nordieik,G.,

Hornischer,K., Brandt,P., Scharte,M., Schoen,O., Desario,A.,

Reichelt,J., Kauer,G., Bloecker,H., Ramser,J., Beck,A., Klages,S.,

Hennig,S., Rieselmann,L., Dagand,E., Haaf,T., Wehrmeyer,S.,

Borzym,K., Gardiner,K., Nizetic,D., Francis,F., Lehrach,H.,

Reinhardt,R. and Yaspo,M.Laure.

The DNA sequence of human chromosome 21

Unpublished

REFERENCE

2 (bases 1 to 190837)

Taudien,S., Dagand,E., Delabar,J., Orti,R., Nordieik,G.,

Drescher,B., Weber,J., Schattevoy,R., Menzel,U., Yaspo,M.-L. and

Rosenthal,A.

Direct Submission

Submitted (16-FEB-1999) Genome Analysis, Institute of Molecular

Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany

REFERENCE

3 (bases 1 to 190837)

Taudien,S., Dagand,E., Delabar,J., Orti,R., Nordieik,G.,

Drescher,B., Weber,J., Schattevoy,R., Menzel,U., Yaspo,M.-L. and

Rosenthal,A.

Direct Submission

Submitted (20-APR-2001) Genome Analysis, Institute of Molecular

Biotechnology, Beutenbergstr. 11, Jena 07745, Germany

REFERENCE

4 (bases 1 to 190837)

Legemann,D. and Platzer,M.

Direct Submission

Submitted (13-AUG-2003) Genome Analysis, Institute of Molecular

Biotechnology, Beutenbergstr. 11, Jena 07745, Germany

TITLE

JOURNAL

COMMENT

On Aug 13, 2003 this sequence version replaced gi:13699324.

Center: Genome Center

Center: Institute of Molecular Biotechnology

Center code: IMB

Web site: <http://genome.imb-jena.de/>Contact: gscj-submit@genome.imb-jena.de

Project Information

Center project name: U70-64

Center clone name: RP1-100J12, RP1-P79E4

Consensus quality: 188276 bases at least Q40

Consensus quality: 130662 bases at least Q30

Consensus quality: 190837 bases at least Q20

Quality coverage: 10.52x

This sequence was finished as follows unless otherwise noted: all

regions were double stranded, sequenced with an alternate

chemistry, or covered by high quality data (i.e., phred quality >=

30); an attempt was made to resolve all sequencing problems, such

as compressions and repeats; all regions were covered by at least

one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest.

Neighboring sequence information:

This clone is overlapped by RP5-866H8, RP1-84N21.

FEATURES

source

1. .190837

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="21"

/map="q21.3-22.11"

/clone="RP1-100J12, RP1-P79E4"

1. .121509

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/clone="RP1-100J12"

48002. .190837

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/clone="RP1-79E4"

653

/note="C substituted in clone: RP5-866H8"

/replaces="t"

1060

/note="A substituted in clone: RP5-866H8"

/replaces="g"

1314

/note="C substituted in clone: RP5-866H8"

/replaces="t"

3311

/note="C substituted in clone: RP5-866H8"

/replaces="t"

3725

/note="G substituted in clone: RP5-866H8"

/replaces="a"

3827

/note="deleted in clone: RP1-100J12"

/replaces=""

5555

/note="T substituted in clone: RP5-866H8"

/replaces="c"

6971

/note="deleted in clone: RP5-866H8"

/replaces=""

7254

/note="G substituted in clone: RP5-866H8"

/replaces="c"

8668

/note="G substituted in clone: RP5-866H8"

/replaces="a"

15237. .15255

/note="single stranded/single chemistry region"

26658. .26695

/note="single stranded/single chemistry region"

30914. .31021

/note="single stranded/single chemistry region"

30916. .30919

/note="low quality region"

30919

39287. .39356

/note="single stranded/single chemistry region"

39705. .39766

/note="single stranded/single chemistry region"

44251

/note="low quality region"

44287. .44367

/note="single stranded/single chemistry region"

48080

/note="C substituted in clone: RP1-79E4"

```

/replace="g"
45109 /note="G substituted in clone: RP1-79E4"
/replace="t"
45195 /note="C substituted in clone: RP1-79E4"
/replace="t"
45693 /note="G substituted in clone: RP1-100J12"
/replace="a"
52403 /note="C substituted in clone: RP1-79E4"
/replace="a"
53247..53248 /note="C substituted in clone: RP1-79E4"
/replace="a"
53545 /note="deleted in clone: RP1-79E4"
/replace="a"
53696 /note="A substituted in clone: RP1-79E4"
/replace="c"
58070 /note="C substituted in clone: RP1-79E4"
/replace="t"
60115 /note="deleted in clone: RP1-79E4"
/replace="a"
61768..61852 /note="single stranded/single chemistry region"
63716..63717 /note="deleted in clone: RP1-100J12"
/replace="a"
70290..70297 /note="deleted in clone: RP1-79E4"
/replace="a"
73369..73571 /note="pcr product sequence only"
73856..74096 /note="pcr product sequence only"
74003..74125 /note="single stranded/single chemistry region"
74435 /note="T substituted in clone: RP1-79E4"
/replace="a"
74449 /note="C substituted in clone: RP1-79E4"
/replace="g"
74490 /note="A substituted in clone: RP1-79E4"
/replace="g"
75076 /note="A substituted in clone: RP1-79E4"
/replace="g"
75753 /note="deleted in clone: RP1-100J12"
/replace="a"
78907 /note="G substituted in clone: RP1-79E4"
/replace="a"
79856..79883 /note="single stranded/single chemistry region"

Query Match 79.2%; Score 19.8; DB 8; Length 190837;
Best Local Similarity 91.3%; Pred. No. 90;
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 AAAAAACACTTGGAGGAGACT 23
||||| ||||| ||||| |||||
Db 113272 AAAAAAACTTGTAAAGGAGACT 113250

```

RESULT 22

```

BS000157/c
LOCUS      192836 bp      DNA      linear      PRI 09-JUN-2004
DEFINITION Pan troglodytes chromosome 22 clone:RP43-154K22, map 22, complete
            sequences.
ACCESSION  BS000157
VERSION    BS000157.1
KEYWORDS   GI:37537424
SOURCE     Pan troglodytes (chimpanzee)
ORGANISM   Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pan.
REFERENCE  1
            The International Chimpanzee Chromosome 22 Consortium.
            DNA sequence and comparative analysis of chimpanzee chromosome 22
            Nature 429, 382-388 (2004)
            2 (bases 1 to 192836)
            Hattori,M., Toyoda,A., Watanabe,H., Taylor,T.D., Kuroki,Y.,
            Fujiyama,A. and Sakaki,Y.
            Direct Submission
            Submitted (12-MAY-2003) Masahira Hattori, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC); Japan
            1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            (E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
            Tel:81-45-503-9111, Fax:81-45-503-9170)
            The Chimpanzee Chromosome 22 Sequencing Consortium consists of:
            *Chinese National Human Genome Center at Shanghai, Shanghai, China;
            *GBF, Dept. of Genome Analysis, Braunschweig, Germany; *Institute
            of Molecular Biotechnology, Jena, Germany; *KRIBB Genome Research
            Center, Daejeon, Korea;
            *Max-Planck-Institute for Molecular Genetics, Berlin, Germany;
            *National Institute of Genetics, Mishima, Japan;
            *National Yang Ming University Genome Research Center, Taipei,
            Taiwan;
            *RIKEN Genomic Sciences Center, Yokohama, Japan.
            ----- Genomic Center
            Center: RIKEN Genomic Sciences Center
            Center code: RIKEN
            Web site: http://hgp.gsc.riken.go.jp/
            Contact: hattori@gsc.riken.go.jp
            ----- Project Information
            Center project name:The Chimpanzee Chromosome 22 Sequencing Project
            Center Clone name: RP43-154K22
            ----- Summary Statistics
            Sequencing vector: pUC18,pUC13,prz19R; 100% of reads Chemistry:
            Dye-terminator Big Dye and ET; 100% of reads Assembly Program:
            Phrap; version 0.990329
            Consensus quality: 191,815 bases at least Q40
            Consensus quality: 960 bases at least Q30
            Consensus quality: 61 bases at least Q20
            -----
            This sequence was finished as follows unless otherwise noted: all
            regions were double stranded, sequenced with an alternate
            chemistry, or covered by high quality data (i.e., phred quality >=
            30);
            an attempt was made to resolve all sequencing problems, such as
            compressions and repeats; all regions were covered by at one
            plasmid
            subclone or more than one M13 subclone;
            and the assembly was confirmed by restriction digest.
            -----
            Source information:
            The RPCI-43 chimpanzee BAC library was prepared from DNA isolated
            from the blood of a single male chimpanzee using published
            protocols (Osoegawa,K. et al. Genomics 52:1-8). The DNA from the
            chimpanzee ('Clint') was obtained from the Yerkes Primate Center in
            Atlanta. The library was prepared by Baoli Zhu, Chung Li Shu,
            Kazutoyo Osoegawa, Evan Eichler & Pieter J de Jong. The library
            characteristics are described at
            http://www.chori.org/bacpac/mchimp43.htm.
            The clone may be obtained from Pieter J. de Jong and coworkers
            (http://www.chori.org/bacpac).
            VECTOR: pBACe3.6

```

The CHORI-251 chimpanzee BAC library was prepared from DNA isolated from the blood of a single male chimpanzee using published protocols (Osoegawa, K. et al. Genomics 52:1-8). The DNA from the chimpanzee ('Cline') was obtained from the Yerkes Primate Center in Atlanta. The library was prepared by Baoji Zhu, Chung Li Shu, Kazutoyo Osoegawa, Evan Bichler & Pieter J de Jong. The library characteristics are described at <http://www.chori.org/bacpac/chimpanzee251.htm>. The clone may be obtained from Pieter J. de Jong and coworkers (<http://www.chori.org/bacpac>).

VECTOR: pTARBAC2.1
The PTB1 chimpanzee BAC library was prepared from DNA isolated from cultured cells established from the blood of a single male chimpanzee.

Clones may be obtained from Asao Fujiyama and co-workers (<http://www.gsc.riken.go.jp>).

VECTOR: pKS145

The PTB22 chimpanzee Fosmid library was prepared from DNA isolated from cultured cells established from the blood of a single male chimpanzee.

Clones may be obtained from Asao Fujiyama and co-workers (<http://www.gsc.riken.go.jp>).

VECTOR: pKS143

Sequence Quality Assessment:

This entry has been annotated with sequence

estimates computed by the Phrap assembly program.

All manually edited bases have been reduced to quality zero.

Quality levels above 40 are expected to have less than 1 error in 10,000 bp.

Neighboring clones: PTB-083K24 (left) and PTB-086M04 (right).

FEATURES

source
Location/Qualifiers
1..192836
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/chromosome="22"
/clone="RP43-154K22"
/clone_lib="RPCI-43 chimpanzee BAC"

ORIGIN

Query Match 79.2%; Score 19.8; DB 8; Length 192836;
Best Local Similarity 91.3%; Pred. No. 89;
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAACACTTGGGAAGGAGACT 23
||||| ||||| ||||| ||||| |||||
Db 25268 AAAAAACACTTGGGAAGGAGACT 25246

RESULT 23
AL954132/c 197745 bp DNA linear VRT 05-JUN-2003
LOCUS Zebrafish DNA sequence from clone DKEY-32N7, complete sequence.
DEFINITION
ACCESSION AL954132
VERSION AL954132.14 GI:31455429
KEYWORDS HTG.
SOURCE Danio rerio (zebrafish)

ORGANISM

Danio rerio
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 197745)
Phillimore, B.
Direct Submission
Submitted (05-JUN-2003) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Jun 6, 2003 this sequence version replaced gi:31407707.

REFERENCE

AUTHORS
TITLE
JOURNAL
COMMENT
Center: Wellcome Trust Sanger Institute
Center code: SC

Web site: <http://www.sanger.ac.uk>
Contact: zfish-help@sanger.ac.uk

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP database can be found at

http://www.sanger.ac.uk/Projects/C_elegans/wormpep Repeat names beginning 'Dr' were identified by the Recon repeat discovery system (Zhirong Bao and Sean Eddy, submitted), and those beginning 'drr' were identified by Rick Waterman (Stephen Johnson lab, WashU). For further information see

http://www.sanger.ac.uk/Projects/D_reio/fishmask.shtml DKEY-32N7 is from a Zebrafish BAC library

FEATURES

Location/Qualifiers
1..197745
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="DKEY-32N7"
/clone_lib="DanioKey"

ORIGIN

Query Match 79.2%; Score 19.8; DB 5; Length 197745;
Best Local Similarity 91.3%; Pred. No. 89;
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AAAAAACACTTGGGAAGGAGACT 24
||||| ||||| ||||| ||||| |||||
Db 62477 AAAAAACACTTGGGAAGGAGACT 62455

RESULT 24
BS000171/c 218680 bp DNA linear PRI 12-JUN-2004
LOCUS Pan troglodytes chromosome 22 clone:PTB-086M04, map 22, complete sequences.

ACCESSION BS000171 BA000046
VERSION BS000171.1 GI:37537438
KEYWORDS HTG.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Pan.

REFERENCE
1
The International Chimpanzee Chromosome 22 Consortium.
DNA sequence and comparative analysis of chimpanzee chromosome 22
Nature 429, 382-388 (2004)

REFERENCE
2 (bases 1 to 218680)
Wang, S., Cai, Z., Wang, B., Zheng, H., Zhang, Y., Zhang, X., Zhu, G.,
Lu, G., Fu, G. and Chen, Z.

Direct Submission
Submitted (26-MAY-2003) Shengyue Wang, Chinese National Human
Genome Center at Shanghai, Shanghai 201203, CHINA
Zhang Jiang HI-TECH Park, Shanghai 201203, CHINA
(E-mail: wangyue@chgc.sh.cn, URL: <http://www.chgc.sh.cn>,
Tel: 86-21-50801919, Fax: 86-21-50801922)

JOURNAL

Submitted (30-MAY-2005) Human Genome Sequencing Center, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 On Jan 29, 2005 this sequence version replaced gi:56676412.
 Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality does not meet this standard, it will be indicated in the annotation.

COMMENT

The repeat regions shown were identified using RepeatMasker by Adrian Smit.

Sequence similarities were identified using Powerblast by Jinghui Zhang.

Exon/Intron boundaries of identified genes were chosen if there were canonical splice junctions that maintained sequence continuity across the splice junctions.

FEATURES

source
 1. 219322
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /chromosome="6"
 /clone="RP23-13B24"
 complement(1..142)
 /rpt_family="L1_Mm"
 2059..2244
 /rpt_family="Lx4B"
 2295..2389
 /rpt_family="GA"
 2423..2515
 /rpt_family="GA"
 3746..3803
 /rpt_family="TTTG"
 3815..3846
 /rpt_family="CCA"
 complement(4969..5013)
 /rpt_family="L3"
 5632..5680
 /rpt_family="TTA"
 complement(5771..6024)
 /rpt_family="Lx8"
 7935..8304
 /rpt_family="L1Md_F3"
 complement(8335..8699)
 /rpt_family="ORR1D"
 complement(8940..9174)
 /rpt_family="L1_Mur3"
 complement(9175..9641)
 /rpt_family="L1_Mus3"
 complement(9642..9674)
 /rpt_family="L1_Mur3"
 9900..10071
 /rpt_family="L1Md_F2"
 12020..12046
 /rpt_family="TTTG"
 17140..17222
 /rpt_family="CA"
 17277..17382
 /rpt_family="MLTB"
 complement(17541..17588)
 /rpt_family="L2"
 17949..18012
 /rpt_family="CT-rich"
 18014..18124
 /rpt_family="C-rich"
 18132..18270
 /rpt_family="CT-rich"
 complement(19103..19270)
 /rpt_family="L1Md_F2"
 22045..22786
 /rpt_family="L1Md_F2"

repeat_region
 /rpt_family="T"
 complement(25197..25323)
 /rpt_family="B1_Mus1"
 27714..27723
 /note="single subclone"
 /function="single clone coverage"
 29432..29742
 /rpt_family="L1Md_T"
 complement(32227..33010)
 /rpt_family="L1Md_T"
 33234..33355
 /rpt_family="TTTC"
 33385..33528
 /rpt_family="RSINE1"
 34120..34221
 /rpt_family="RSINE1"
 34247..34274
 /rpt_family="TAAAAA"
 34887..35090
 /rpt_family="B3"
 35546..35567
 /rpt_family="AT-rich"
 complement(36257..36351)
 /rpt_family="MTE"
 37477..37918
 /rpt_family="L1Md_F3"
 complement(37939..38080)
 /rpt_family="B1_Mus1"
 38397..38447
 /rpt_family="CA"
 40153..40204
 /rpt_family="ATG"
 complement(40354..40521)
 /rpt_family="URR1A"
 complement(40872..41236)
 /rpt_family="ORR1D"
 41697..41726
 /rpt_family="AT-rich"
 complement(42404..42502)
 /rpt_family="L3"
 42552..42757
 /rpt_family="B4"
 complement(42758..42903)
 /rpt_family="B1_Mm"
 42904..42975
 /rpt_family="B4"
 43338..43360
 /rpt_family="AT-rich"
 44681..44722
 /rpt_family="TG"
 45747..45804
 /rpt_family="GA-rich"
 45989..46018
 repeat_region

Query Match 79.2%; Score 19.8; DB 9; Length 219322;
 Best Local Similarity 91.3%; Pred. No. 88;
 Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 AAAACACTTGGAAAGGAGACTCA 25
 |||||

Db 208002 AAAAGCTTGGAAAGGAGACTCA 207980

RESULT 26

HS21C049/c
 LOCUS HS21C049 34000 bp DNA linear
 DEFINITION Homo sapiens chromosome 21 segment HS21C049.
 ACCESSION AL163249 AP001704 BA000005
 VERSION AL163249.2 GI:7717307
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 340000)

REFERENCE AUTHORS

Hattori,M., Fujiyama,A., Taylor,T.D., Watanabe,H., Yada,T., Park,H.S., Toyoda,A., Iehii,K., Totoki,Y., Choi,D.K., Soeda,E., Ohki,M., Takagi,T., Sakaki,Y., Taudien,S., Blechschmidt,K., Polley,A., Menzel,U., Delabar,J., Kumpf,K., Lehmann,R., Patterson,D., Reichwald,K., Rump,A., Schillhabel,M., Schudy,A., Zimmermann,W., Rentschel,A., Kudoh,J., Shibuya,K., Kawasaki,K., Asakawa,S., Shintani,A., Sasaki,T., Nagamine,K., Mitsuyama,S., Antonarakis,S.E., Minoshima,S., Shimizu,N., Nordseick,G., Hornischer,K., Brandt,P., Scharfe,M., Schoen,O., Desario,A., Reichelt,J., Kauer,G., Bloecker,H., Ramser,J., Beck,A., Klages,S., Hennig,S., Riesselmann,L., Dagand,E., Wehrmeyer,S., Borzym,K., and Gardiner,K., Nizetic,D., Francis,F., Lehrach,H., Reinhardt,R., and Yaspo,M.L.

TITLE JOURNAL

Submitted (05-MAY-2000) The Chromosome 21 Mapping and Sequencing Consortium: * RIKEN Genomic Sciences Center, Human Genome Research Group * Institute of Molecular Biotechnology, Genome Analysis * Keio University School of Medicine, Dept. of Molecular Biology * GBF, Dept. of Genome Analysis * Max-Planck Institute for Molecular Genetics (addresses see below)

The Chromosome 21 Mapping and Sequencing Consortium consists of

* RIKEN Genomic Sciences Center, Human Genome Research Group, * Sagamihara 228-8535, Japan,
* e.mail: sakaki@gsc.riken.go.jp
* URL: http://hgp.gsc.riken.go.jp/

and

* Institute of Molecular Biotechnology, Genome Analysis, * Beutenbergstrasse 11, D-07745 Jena, Germany,
* e.mail: gscj-submit@genome.imb-jena.de
* URL: http://genome.imb-jena.de/

and

* Keio University School of Medicine, Dept. of Molecular Biology, * Tokyo 160-8582 Japan,
* e.mail: shimizu@mb-med.keio.ac.jp
* URL: http://adenine.dmb.med.keio.ac.jp/

and

* GBF, Dept. of Genome Analysis,
* Mascheroder Weg 1, D-38124 Braunschweig, Germany, * e.mail: info.genome@gbf.de

and

* Max-Planck Institute for Molecular Genetics,
* Ihnestrasse 73, D-14195 Berlin, Germany,
* e.mail: info-chr21@molgen.mpg.de
* URL: http://chr21.rz-berlin.mpg.de/.

FEATURES

source

1..340000
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.1"
<1..10570
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.1"
/clone="P866H8, 5' partial"
/clone_lib="RPC11,3-5 PAC library"
1590..121668
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.1"
/clone="P100J12"
/clone_lib="RPC11,3-5 PAC library"
/note="Accession No. AF129075"

source

complement (1930..2178)
/gene="ZNF294"
/number=4
2434..2737
/note="AluSx"
/rpt_family="SINE/Alu"
/rpt_type=DISPERSED
complement (2915..3216)
/note="AluSx"
/rpt_family="SINE/Alu"
/rpt_type=DISPERSED
complement (3395..3493)
/gene="ZNF294"
/number=3
complement (3987..4190)

source

49591..190996
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.1"
/clone="P795E4"
/clone_lib="RPC11,3-5 PAC library"
/note="Accession No. AF129075"
169378..274529
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.1"
/clone="P84N21"
/clone_lib="RPC11,3-5 PAC library"
/note="Accession No. AF124730"
258688..>340000
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.1"
/clone="P292A20, 3' partial"
/clone_lib="RPC11,3-5 PAC library"
/note="Accession No. AF124731"
290605..>340000
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.1"
/clone="P71073, 3' partial"
/clone_lib="CMP21 P1 library"
/note="Accession No. AP000241"
complement (1930..>10199)
/gene="ZNF294"
complement (join (1930..2178,3395..3493,3987..4190,10020..>10199))
/gene="ZNF294"
complement (join (1930..2178,3395..3493,3987..4190,10020..10199))
/note="mRNA consists of Accession Nos. AB018257 and AF150217, junction confirmed by RT-PCR"
/codon_start=1
/product="containing human KIAA0714 protein"
/protein_id="CAB90430.1"
/db_xref="GI:7717308"
/db_xref="Gene:13082"
/db_xref="GOA:O94822"
/db_xref="UniProt/Swiss-Prot:O94822"
/translation="MDRVGGKGPGSYRVSSSSAARRKLCQGRVNRGLSAQSTAAT
ATMGKNKQRTKGNLRPSNGRAELLAKEQTVPGFIGTSGDGLGVVPAIQAAE
IDSLVDSDFRMVRLKLSKQVTKLKAQEFQCTERTETVKGLVPWPRI FCKIS
LDHDRVREATQQAPEKLIKVKQLAPFLKSLMGYWLMAQCDDTTTPAFAAKDAFEA
APFSPQPEAIAPFCKDEITSVSSGII"
complement (1930..2178)
/gene="ZNF294"
/number=4
2434..2737
/note="AluSx"
/rpt_family="SINE/Alu"
/rpt_type=DISPERSED
complement (2915..3216)
/note="AluSx"
/rpt_family="SINE/Alu"
/rpt_type=DISPERSED
complement (3395..3493)
/gene="ZNF294"
/number=3
complement (3987..4190)

* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 702: contig of 702 bp in length
* 703 782: gap of unknown length
* 783 1322: contig of 540 bp in length
* 1323 1402: gap of unknown length
* 1403 84895: contig of 83493 bp in length
* 84896 84975: gap of unknown length
* 84976 85473: contig of 498 bp in length
* 85474 85553: gap of unknown length
* 85554 86109: contig of 556 bp in length
* 86110 86189: gap of unknown length
* 86190 86795: contig of 606 bp in length
* 86796 86875: gap of unknown length
* 86876 87456: contig of 581 bp in length
* 87457 87536: gap of unknown length
* 87537 88144: contig of 608 bp in length
* 88145 88224: gap of unknown length
* 88225 88871: contig of 647 bp in length
* 88872 88951: gap of unknown length
* 88952 89406: contig of 455 bp in length
* 89407 89486: gap of unknown length
* 89487 90083: contig of 597 bp in length
* 90084 90163: gap of unknown length
* 90164 90788: contig of 625 bp in length.

FEATURES

source

Location/Qualifiers
1..90788
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/strain="y; cn bw sp"
/db_xref="taxon:7227"
/chromosomes="2"
/map="36B3-36B4"
/clone="DS07161 (D356)"
/clone_lib="P1 library, partial Sau3A in pAd10sacBII"
703..782
/estimated_length=unknown
1323..1402
/estimated_length=unknown
84896..84975
/estimated_length=unknown
85474..85553
/estimated_length=unknown
86110..86189
/estimated_length=unknown
86796..86875
/estimated_length=unknown
87457..87536
/estimated_length=unknown
88145..88224
/estimated_length=unknown
88872..88951
/estimated_length=unknown
89407..89486
/estimated_length=unknown
90084..90163
/estimated_length=unknown

ORIGIN

Query Match 77.6%; Score 19.4; DB 14; Length 90788;
Best Local Similarity 95.2%; Pred. No. 1.5e+02;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AAAAAACACTTGAAGGGAGA 21

Db 73424 AAAAAACACTTGAAGGGAGA 73444

RESULT 29

AC009250/c 170362 bp DNA linear INV 10-MAR-2001
LOCUS Drosophila melanogaster, Chromosome 2L, region 36B-36B, BAC clone

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS

TITLE
JOURNAL
REFERENCE
AUTHORS

TITLE
JOURNAL
COMMENT

FEATURES
source

ORIGIN

Query Match 77.6%; Score 19.4; DB 2; Length 170362;
Best Local Similarity 95.2%; Pred. No. 1.4e+02;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AAAAAACACTTGAAGGGAGA 21

Db 2196 AAAAAACACTTGAAGGGAGA 2176

RESULT 30

BACR27M18, complete sequence.

AC009250 GI:13270514

HTG.

Drosophila melanogaster (fruit fly)

Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.

1 (bases 1 to 170362)

Celniker, S.E., Adams, M.D., Kronmiller, B., Tyler, D., Wan, K.H.,
Holt, R.A., Evans, C.A., Gocayne, J.D., Amanatides, P.G., Brandt, R.C.,
Rogers, Y., Han, H., Baldwin, D., Banzon, J., Beeson, K.Y., Busam, D.A.,
Carlson, J.W., Center, A., Chao, M., Davenport, L.B., Dietz, S.M.,
Dodson, K., Dorsett, V., Doup, L.E., Doyle, C., Dresnek, D., Farfan, D.,
Ferrera, S., Frise, E., Galle, R.F., Garg, N.S., George, R.A.,
Gonzalez, M., Houck, J., Hoskins, R.A., Hostin, D., Howland, T.J.,
Ibegwam, C., Jalali, M., Kruse, D., Li, P., Mattei, B., Moshrefi, A.,
McIntosh, T.C., Moy, M., Murphy, B., Nelson, C., Nelson, K.A., Nunoo, J.,
Paclet, J., Paragas, V., Park, S., Patel, S., Pfeiffer, B.,
Phouanavong, S., Pittman, G.S., Puri, V., Richards, S., Scheeler, F.,
Stapleton, M., Strong, R., Svirskaas, R., Tector, C., Williams, S.M.,
Zaveri, J.S., Smith, H.O., Rubin, G.M. and Venter, J.C.
Sequencing of Drosophila chromosome 2L, region 36B-36B

Unpublished

2 (bases 1 to 170362)

Celniker, S.E., Agbayani, A., Arcaina, T.T., Baxter, E., Blazej, R.G.,
Butenhoff, C., Champe, M., Chavez, C., Chew, M., Ciesiolka, L.,
Doyle, C.M., Farfan, D.E., Galle, R., George, R.A., Harris, N.L.,
Hoskins, R.A., Houston, K.A., Hummasti, S.R., Karra, K., Kearney, L.,
Kim, E., Lee, B., Lewis, S., Li, P., Lomotan, M.A., Mazda, P.,
Moshrefi, A.R., Moshrefi, M., Nixon, K., Paclet, J.M., Park, S.,
Pfeiffer, B., Poon, L., Sequeira, A., Sethi, H., Snir, E.,
Svirskaas, R.R., Wan, K.H., Weinburg, T., Zhang, R., Zieran, L.L. and
Rubin, G.M.

Direct Submission

Submitted (09-AUG-1999) Drosophila Genome Center, Lawrence Berkeley
Laboratory, MS 64-121, Berkeley, CA 94720, USA

On Mar 10, 2001 this sequence version replaced gi:5734689.

Sequence submitted by:

Lawrence Berkeley National Laboratory

Berkeley, CA 94720

This sequence was assembled using end sequences from a whole genome
shotgun and from subclones of this BAC and its neighboring clones.
For further information about this sequence, including its location
and relationship to other sequences, please visit our sequence
archive Web site (<http://www.fruitfly.org/sequence/>) or send email
to bdgp@fruitfly.berkeley.edu.

Location/Qualifiers

1..170362

/organism="Drosophila melanogaster"

/mol_type="genomic DNA"

/strain="y; cn bw sp"

/db_xref="taxon:7227"

/chromosome="2L"

/map="36B-36B"

/clone="BACR27M18 (D1027)"

/clone_lib="RPCI-98 (Roswell Park Cancer Institute
Drosophila melanogaster BAC library, partial EcoRI in
pBACe3.6)"

AC090872	AC090872	178481 bp	DNA	linear	INV 14-MAR-2001
LOCUS	Drosophila melanogaster, chromosome 2L, region 36B-36C, BAC clone				
DEFINITION	BACR06G19, complete sequence.				
ACCESSION	AC090872	GI:13324747			
VERSION	AC090872.1				
KEYWORDS	HTG.				
SOURCE	Drosophila melanogaster (fruit fly)				
ORGANISM	Drosophila melanogaster				
REFERENCE	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.				
AUTHORS	1 (bases 1 to 178481) Celniker,S.E., Adams,M.D., Krommiller,B., Tyler,D., Wan,K.H., Holt,R.A., Evans,C.A., Gocayne,J.D., Amanatides,P.G., Brandon,R.C., Rogers,Y., An,H., Baldwin,D., Banazon,J., Beeson,K.Y., Busam,D.A., Carlson,J.W., Center,A., Champe,M., Davenport,L.B., Dietz,S.M., Dodson,K., Dorsett,V., Doup,L.E., Doyle,C., Dresnek,D., Farfan,D., Ferreira,S., Frise,E., Galle,R.F., Garg,N.S., George,R.A., Gonzalez,M., Houck,J., Hoskins,R.A., Hostin,D., Howland,T.J., Ibegwam,C., Jalali,M., Kruse,D., Li,P., Mattei,B., Moshrefi,A., McIntosh,T.C., Moy,M., Murphy,B., Nelson,C., Nelson,K.A., Nunoo,J., Pacleb,J., Paragas,V., Park,S., Patel,S., Pfeiffer,B., Phouanenavong,S., Pittman,G.S., Puri,V., Richards,S., Scheeler,F., Stapleton,M., Strong,R., Svirskaas,R., Tector,C., Williams,S.M., Zaveri,J.S., Smith,H.O., Rubin,G.M. and Venter,J.C.				
TITLE	Sequencing of Drosophila chromosome 2L, region 36B-36C				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 178481) Celniker,S.E., Adams,M.D., Krommiller,B., Tyler,D., Wan,K.H., Holt,R.A., Evans,C.A., Gocayne,J.D., Amanatides,P.G., Brandon,R.C., Rogers,Y., An,H., Baldwin,D., Banazon,J., Beeson,K.Y., Busam,D.A., Carlson,J.W., Center,A., Champe,M., Davenport,L.B., Dietz,S.M., Dodson,K., Dorsett,V., Doup,L.E., Doyle,C., Dresnek,D., Farfan,D., Ferreira,S., Frise,E., Galle,R.F., Garg,N.S., George,R.A., Gonzalez,M., Houck,J., Hoskins,R.A., Hostin,D., Howland,T.J., Ibegwam,C., Jalali,M., Kruse,D., Li,P., Mattei,B., Moshrefi,A., McIntosh,T.C., Moy,M., Murphy,B., Nelson,C., Nelson,K.A., Nunoo,J., Pacleb,J., Paragas,V., Park,S., Patel,S., Pfeiffer,B., Phouanenavong,S., Pittman,G.S., Puri,V., Richards,S., Scheeler,F., Stapleton,M., Strong,R., Svirskaas,R., Tector,C., Williams,S.M., Zaveri,J.S., Smith,H.O., Rubin,G.M. and Venter,J.C.				
TITLE	Direct Submission				
JOURNAL	Submitted (14-MAR-2001) Berkeley Drosophila Genome Project, MS 64-121, Lawrence Berkeley National Laboratory, One Cyclotron Road, Berkeley, CA 94720, US				
COMMENT	Sequence submitted by: Lawrence Berkeley National Laboratory, MS 64-121 Berkeley, CA 94720				
FEATURES	This sequence was assembled using end sequences from a whole genome shotgun and from subclones of this BAC and its neighboring clones. For further information about this sequence, including its location and relationship to other sequences, please visit our sequence archive Web site (http://www.fruitfly.org/sequence/) or send email to bdgp@fruitfly.berkeley.edu . Location/Qualifiers 1. 178481 /organism="Drosophila melanogaster" /mol_type="genomic DNA" /strain="y; cn bw sp" /db_xref="taxon:7227" /chromosome="2L" /map="36B-36C" /clone="BACR06G19 (DI1335)" /clone_lib="RPCI-98 (Roswell Park Cancer Institute Drosophila melanogaster BAC library, partial EcoRI in pBACe3.6)"				
ORIGIN	Query Match 77.6%; Score 19.4; DB 2; Length 178481; Best Local Similarity 95.2%; Pred. No. 1.4e+02; Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;				
TITLE	Direct Submission				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 202907) Worley,K.C.				
AUTHORS					
QY	1	AAAAAACACTTGGAGGGGAGA	21		
DB	27034	AAAAACACTTGGATGGGAGA	27054		
RESULT 31	AC152322	202907 bp	DNA	linear	HTG 01-JUL-2005
LOCUS	Bos taurus clone CH240-207, WORKING DRAFT SEQUENCE, 13 unordered pieces.				
DEFINITION	AC152322				
ACCESSION	AC152322.3	GI:69227244			
VERSION	HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.				
KEYWORDS	Bos taurus (cow)				
SOURCE	Bos taurus				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.				
REFERENCE	1 (bases 1 to 202907)				
AUTHORS	Muzny,D.Marie., Metzker,M.Lee., Abramson,S., Adams,C., Alder,J., Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D., Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H., Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F., Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M., Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E., Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J., Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L., Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D., Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K., Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K., Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G., Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P., Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M., Gebregorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W., Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K., Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J., Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hoques,M., Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A., Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A., Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C., Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J., Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J., Lorenshewa,L., Loulseghe,H., Lozado,R.J., Lu,X., Ma,J., Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A., Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E., Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E., Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S., Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L., Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Nwankwelen,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K., Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkuch,C., Plopper,F., Polindexter,A., Popovic,D., Primus,E., Pu,L., L., Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F., Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J., Sanders,W., Savary,G., Schefer,S., Scott,G., Shatsman,S., Shen,H., Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajz,D., Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J., Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K., Valas,R., Vera,V., Villaseana,D., Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F., Williams,G., Willson,R., Wlaczuk,R., Woodden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,X., Zhou,X., Zhao,S., Dunn,D., von Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O., Weinstock,G. and Gibbs,R.A.				
TITLE	Direct Submission				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 202907)				
AUTHORS	Worley,K.C.				

TITLE
JOURNAL

REFERENCE
AUTHORS

TITLE
JOURNAL

COMMENT

Direct Submission
Submitted (01-NOV-2004) Human Genome Sequencing Center, Department
Of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 202907)
Cow Genome Sequencing Consortium.
Direct Submission
Submitted (01-JUN-2005) Human Genome Sequencing Center, Department
Of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jun 26, 2005 this sequence version replaced gi:58038140.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: FARP
Center clone name: CH240-207
----- Summary Statistics
Assembly program: Atlas 3.0;
Consensus quality: 199720 bases at least Q40
Consensus quality: 200773 bases at least Q30
Consensus quality: 201499 bases at least Q20
Estimated insert size: 203751; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 13 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved

* 34144: contig of 34144 bp in length
* 34145
* 34194: gap of 50 bp
* 41485: contig of 7291 bp in length
* 41486
* 41535: gap of 50 bp
* 5104: contig of 13569 bp in length
* 55104: gap of 50 bp
* 55154: gap of 50 bp
* 55155
* 71707: contig of 16553 bp in length
* 71757: gap of 50 bp
* 71758
* 76733: contig of 4976 bp in length
* 76734
* 76784
* 78913: contig of 2130 bp in length
* 78914
* 78963: gap of 50 bp
* 78964
* 83371: contig of 4408 bp in length
* 83372: gap of 50 bp
* 83422
* 133340: contig of 49919 bp in length
* 133341
* 133440: gap of unknown length
* 133441
* 160625: contig of 27185 bp in length
* 160626
* 160675: gap of 50 bp
* 160676
* 163035: contig of 2360 bp in length
* 163036
* 163085: gap of 50 bp
* 163086
* 187470: contig of 24385 bp in length
* 187471
* 187520: gap of 50 bp
* 201311: contig of 13791 bp in length
* 201312
* 201411: gap of unknown length

* 201412 202907: contig of 1496 bp in length.

FEATURES
source

1. 202907
Location/Qualifiers
/organism="Bos taurus"
/mol_type="genomic DNA"
/db_xref="taxon:9913"
/clone="CH240-207"
34145..34194
/estimated_length=50
41486..41535
/estimated_length=50
55105..55154
/estimated_length=50
71708..71757
/estimated_length=50
76734..76783
/estimated_length=50
78914..78963
/estimated_length=50
83372..83421
/estimated_length=50
133341..133440
/estimated_length=unknown
160626..160675
/estimated_length=50
163036..163085
/estimated_length=50
187471..187520
/estimated_length=50
201312..201411
/estimated_length=unknown

ORIGIN

Query Match 77.6%; Score 19.4; DB 14; Length 202907;
Best Local Similarity 95.2%; Pred. NO. 1.4e-02;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAAAACACTTGGAGGAGGAGA 21
|||||||

Db 121799 AAAAAACACGTGGAGGAGGAGA 121779
|||||||

RESULT 32
AE003653

LOCUS
DEFINITION
Drosophila melanogaster chromosome 2L, section 62 of 83 of the
complete sequence.
ACCESSION
AE003653 AE002690 AE014134
VERSION
AE003653.3 GI:22946685

KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS

Drosophila melanogaster (fruit fly)
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 247634)
Adams,M.D., Celniker,S.E., Holt,R.A., Evans,C.A., Gocayne,J.D.,
Amanatides,P.G., Scherer,S.E., Li,P.W., Hoskins,R.A., Galle,R.F.,
George,R.A., Lewis,S.E., Richards,S., Ashburner,M., Henderson,S.N.,
Sutton,G.G., Wortman,J.R., Yandell,M.D., Zhang,Q., Chen,L.X.,
Brandon,R.C., Rogers,Y.H., Blazer,R.G., Champe,M., Pfeiffer,B.D.,
Wan,K.H., Doyle,C., Baxter,E.G., Helt,G., Nelson,C.R., Gabor,G.L.,
Abell,J.F., Agbayani,A., An,H.J., Andrews-Pfannkoch,C., Baldwin,D.,
Ballwey,R.M., Basu,A., Baxendale,J., Bayraktaroglu,L., Beasley,E.M.,
Beeson,K.Y., Benos,P.V., Berman,B.P., Bhandari,D., Bolshakov,S.,
Borkova,D., Botchan,M.R., Bouck,J., Brokstein,P., Brottier,P.,
Burtis,K.C., Busam,D.A., Butler,H., Cadiou,E., Center,A.,
Chandra,I., Cherry,J.M., Cawley,S., Dahlke,C., Davenport,L.B.,
Davies,P., de Pablos,B., Delcher,A., Deng,Z., Mays,A.D., Dew,I.,
Dietz,S.M., Dodson,K., Doup,L.E., Downes,M., Dugan-Rocha,S.,
Dunkov,B.C., Dunn,P., Durbin,K.J., Evangelista,C.C., Ferraz,C.,
Ferrieri,S., Fleischmann,W., Fosler,C., Gabrielian,A.E., Garg,N.S.,
Gelbart,W.M., Glasser,K., Glodek,A., Gong,F., Gorrell,J.H., Gu,Z.,

Guan, P., Harris, M., Harris, N.L., Harvey, D., Heiman, T.J., Hernandez, J.R., Houck, J., Hostin, D., Houston, K.A., Howland, T.J., Wei, M.H., Ibegwam, C., Jatali, M., Kalush, F., Karpen, G.H., Ke, Z., Kennison, J.A., Ketchum, K.A., Kimmel, B.E., Kodira, C.D., Kraft, C., Kravitz, S., Kuip, D., Lai, Z., Laško, P., Lei, Y., Levitsky, A.A., Li, J., Li, Z., Liang, Y., Lin, X., Liu, X., Mattioli, B., McIntosh, T.C., McLeod, M.P., McPherson, D., Merkulov, G., Milshina, N.V., Mobarry, C., Morris, J., Moshrefi, A., Mount, S.M., Moy, M., Murphy, B., Murphy, L., Muzny, D.M., Nelson, D.L., Nelson, D.R., Nelson, K.A., Nixon, K., Nusskern, D.R., Pacleb, J.M., Palazzolo, M., Pictman, G.S., Pan, S., Pollard, J.D., Puri, V., Reese, M.G., Reinert, K., Remington, K., Saunders, R.D., Scheeler, F., Shen, H., Shue, B.C., Siden-Kiamos, I., Simpson, M., Skupski, M.P., Smith, T., Spier, E., Spradling, A.C., Stapleton, M., Strong, R., Sun, E., Svirskas, R., Tector, C., Turner, R., Venter, E., Wang, A.H., Wang, X., Wang, Z.Y., Wasserman, D.A., Weinstein, G.M., Weissbach, J., Williams, S.M., Woodage, W., Worley, K.C., Wu, D., Yang, S., Yao, Q.A., Ye, J., Yeh, R.F., Zaveri, J.S., Zhan, M., Zhang, G., Zhao, Q., Zheng, L., Zheng, X.H., Zhong, F.N., Zhong, W., Zhou, X., Zhu, S., Zhu, X., Smith, H.O., Gibbs, R.A., Myers, E.W., Rubin, G.M. and Venter, J.C.

The genome sequence of *Drosophila melanogaster*
 Science 287 (5461), 2185-2195 (2000)
 10731132

2 (bases 1 to 247634)
 Calniker, S.E., Wheeler, D.A., Kronmiller, B., Carlson, J.W., Halpern, A., Patel, S., Adams, M., Champe, M., Dugan, S.P., Frise, E., Hodgson, A., George, R.A., Hoskins, R.A., Lavery, T., Muzny, D.M., Nelson, C.R., Pacleb, J.M., Park, S., Pfeiffer, B.D., Richards, S., Sodergren, E.J., Svirskas, R., Tabor, P.E., Wan, K., Stapleton, M., Sutton, G.G., Venter, C., Weinstein, G., Scherer, S.E., Myers, E.W., Gibbs, R.A. and Rubin, G.M.

Finishing a whole-genome shotgun: release 3 of the *Drosophila melanogaster* euchromatic genome sequence
 Genome Biol. 3 (12), RESEARCH0079 (2002)
 12537568

3 (bases 1 to 247634)
 Miya, S., Crosby, M.A., Mungall, C.J., Matthews, B.B., Campbell, K.S., Hradecky, P., Huang, Y., Kaminker, J.S., Millburn, G.H., Prochnik, S.E., Smith, C.D., Tupy, J.L., Whitfield, E.J., Bayraktaroglu, L., Berman, B.P., Bettencourt, B.R., Celniker, S.E., de Grey, A.D., Drysdale, R.A., Harris, N.L., Richter, J., Russo, S., Schroeder, A.J., Shu, S.Q., Stapleton, M., Yamada, C., Ashburner, M., Gelbart, W.M., Rubin, G.M. and Lewis, S.E.

Annotation of the *Drosophila melanogaster* euchromatic genome: a systematic review
 Genome Biol. 3 (12), RESEARCH0083 (2002)
 12537572

4 (bases 1 to 247634)
 Kaminker, J.S., Bergman, C.M., Kronmiller, B., Carlson, J., Svirskas, R., Patel, S., Frise, E., Wheeler, D.A., Lewis, S.E., Rubin, G.M., Ashburner, M. and Celniker, S.E.

The transposable elements of the *Drosophila melanogaster* euchromatin: a genomics perspective
 Genome Biol. 3 (12), RESEARCH0084 (2002)
 12537573

5 (bases 1 to 247634)
 Celniker, S., Carlson, J., Wan, K., Pfeiffer, B., Frise, E., George, R., Hoskins, R., Stapleton, M., Pacleb, J., Park, S., Svirskas, R., Smith, E., Yu, C. and Rubin, G.

Berkeley *Drosophila* Genome Project
Drosophila melanogaster release 4 sequence
 Unpublished

6 (bases 1 to 247634)
 Adams, M.D., Celniker, S.E., Gibbs, R.A., Rubin, G.M. and Venter, C.J.

Direct Submission
 Submitted (21-MAR-2000) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
 7 (bases 1 to 247634)

FlyBase
 Direct Submission
 Submitted (22-JUL-2005) FlyBase, Harvard University, Biological Laboratories, 16 Divinity Ave, Cambridge, MA 02138, USA

On Sep 16, 2002 this sequence version replaced gi:10728835.

COMMENT
 FEATURES
 Location/Qualifiers
 1..247634
 /organism="Drosophila melanogaster"
 /mol_type="genomic DNA"
 /db_xref="taxon:7227"
 /chromosome="2L"
 complement(<260..934)
 /gene="CG32832"
 /locus_tag="Dmel_CG32832"
 /old_locus_tag="CG32832"
 /map="36B3-36B3"
 /db_xref="FLYBASE:FBgn0052832"
 complement(join(<260..535,676..716,770..934))
 /gene="CG32832"
 /locus_tag="Dmel_CG32832"
 /old_locus_tag="CG32832"
 /product="CG32832-RA"
 /db_xref="FLYBASE:FBgn0052832"
 complement(join(260..535,676..716,770..875))
 /gene="CG32832"
 /locus_tag="Dmel_CG32832"
 /old_locus_tag="CG32832"
 /note="CG32832 gene product from transcript CG32832-RA"
 /codon_start=1
 /product="CG32832-PA"
 /protein_id="AA110977.1"
 /db_xref="GI:22946686"
 /db_xref="FLYBASE:FBgn0052832"
 /translation="MSKGTGPLSKLYNITISTIDKFPVGAQVLQWSQAPGPRTVFFWA
 PFKWSLVLGSLDLSRPPANISLNCQCSLAVTGLIWSRYSVITPKNYLLAVNIA
 VFLLQGLVMVKHLWRSENRAVFNHSHYPIKSGDW"
 1176..7982
 /gene="CG31743"
 /locus_tag="Dmel_CG31743"
 /old_locus_tag="CG31743"
 /map="36B3-36B3"
 /db_xref="FLYBASE:FBgn0032618"
 join(1176..1579,2248..2319,2694..3460,6923..7394,
 7455..7982)
 /gene="CG31743"
 /locus_tag="Dmel_CG31743"
 /old_locus_tag="CG31743"
 /product="CG31743-RA"
 /db_xref="FLYBASE:FBgn0032618"
 join(3173..3460,6923..7394,7455..7684)
 /gene="CG31743"
 /locus_tag="Dmel_CG31743"
 /old_locus_tag="CG31743"
 /note="CG31743 gene product from transcript CG31743-RA"
 /codon_start=1
 /product="CG31743-PA"
 /protein_id="AAF53584.2"
 /db_xref="GI:22946687"
 /db_xref="FLYBASE:FBgn0032618"
 /translation="MDQLVLSFCISIGFSLGNGSTGLPFEDGNNNNVKITRRSLR
 TQTNQRQEFMRQCELLGDHTQTLDELSELOMDHIVDKHKLXCYVPKACTNM
 KRYLMLTNKNGTDPQLPGSLHVGMTKLYDLSEABQQQLSDYTRILVRH
 PLERLLSAYNKLEQSPSARYQSVSRGQIVKELRPGASNSLGRGDDVSFGFEIQ
 LVTPELSRANQSDYNEHWEVIAKLCNFCVMKNVGVKDTLLDLSALALYLAGADNLT
 FPTGHPSTSRANLRNYPDLPIGAIRKLYDIYEDDDFLRFDYALDEVLGPEFG"
 complement(8781..9600)
 /gene="CG6115"
 /locus_tag="Dmel_CG6115"
 /old_locus_tag="CG6115"
 /map="36B3-36B3"
 /db_xref="FLYBASE:FBgn0040985"
 complement(join(8781..9336,9452..9600))
 /gene="CG6115"
 /locus_tag="Dmel_CG6115"
 /old_locus_tag="CG6115"
 /product="CG6115-RA"
 /db_xref="FLYBASE:FBgn0040985"

```
CDS
complement(join(9121..9336,9452..9493))
/gene="CG6115"
/locus_tag="Dmel_CG6115"
/old_locus_tag="CG6115"
/note="CG6115 gene product from transcript CG6115-RA"
/codon_start=1
/product="CG6115-PA"
/protein_id="AAF53585.1"
/db_xref="GI:7298358"
/db_xref="FLYBASE:FBgn040985"
/translation="MSQRSKVISLYKHLYLGREYPLNGPKQPKFKQIHDAFMNHKD
EQDPKIVALLAQGRYLAKVEALYSLKRYSKQYSYND"
<10662..>19731
/gene="CG15133"
/locus_tag="Dmel_CG15133"
/old_locus_tag="CG15133"
/map="36B3-36B4"
/db_xref="FLYBASE:FBgn032619"
/join(<10662..10784,11842..12681,13117..13396,13452..16602,
16660..16863,16946..17489,17550..17830,17896..18016,
19147..>19731)
/gene="CG15133"
/locus_tag="Dmel_CG15133"
/old_locus_tag="CG15133"
/product="CG15133-RA"

Query Match      77.6%; Score 19.4; DB 2; Length 247634;
Best Local Similarity 95.2%; Pred. No. 1.3e+02;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAAAACACTTGAAGGGGAGA 21
Db 87164 AAAAAACACTTGGATGGGAGA 87184

RESULT 33
MWTAM1
LOCUS      428 bp mRNA linear ROD 18-APR-2005
DEFINITION Mouse mRNA for gamma-7S nerve growth factor (gamma-NGF) fragment.
ACCESSION X00472
VERSION X00472.1 GI:54260
KEYWORDS complementary DNA; nerve growth factor; serine protease.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 428)
AUTHORS Howles,P.N., Dickinson,D.P., DiCaprio,L.L., Woodworth-Gutai,M. and
Gross,K.W.
TITLE Use of a cDNA recombinant for the gamma-subunit of mouse nerve
growth factor to localize members of this multigene family near the
TAM-1 locus on chromosome 7
JOURNAL Nucleic Acids Res. 12 (6), 2791-2805 (1984)
PUBMED 620835
FEATURES
source
1..428
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
<1..408
/notes="unnamed protein product; gamma-NGF"
/codon_start=1
/protein_id="CAA25154.1"
/db_xref="GI:54261"
/db_xref="GOA:Q62284"
/db_xref="InterPro:IPR001254"
/db_xref="UniProt/TREMBL:Q62284"
/translation="SKPADITDVKPTILPTPEPKLGSTCLASGWSITPTKPFQTTDD
LYCNVLKLPNEPCAKAIETKVTDMLCAGEMEGKDTCKGSGGLICDGVLQGITTS
WGHTPCGEPMFQVYTKLNKFTSWIKDTWAKNP"

CDS
RNTXALRC
LOCUS      689 bp mRNA linear ROD 18-APR-2005
DEFINITION Rat tissue mRNA for kallikrein C-terminal region (RC 3.4.21.8).
ACCESSION X03560
VERSION X03560.1 GI:57370
KEYWORDS kallikrein; protease; serine protease.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Rattus.
```

```
Query Match      76.8%; Score 19.2; DB 9; Length 428;
Best Local Similarity 87.5%; Pred. No. 4e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAAACACTTGAAGGGGAGACTCA 25
Db 238 AAAGACACTTGAAGGGTGACTCA 261

RESULT 34
BV161004
LOCUS      627 bp DNA linear STS 15-MAY-2004
DEFINITION RPAMMSQ0038539 Roche Palo Alto Mus musculus STS genomic, sequence
tagged site.
ACCESSION BV161004
VERSION BV161004.1 GI:47264404
KEYWORDS STS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 627)
AUTHORS Usuka,J., Liao,G., Cheng,J., Nguyen,A., Bach,C., Puech,A.,
McPherson,J.D., Foernzler,D. and Peltz,G.
TITLE Mus musculus SNPs
JOURNAL Unpublished (2003)
COMMENT Contact: Jonathan Usuka
Roche Palo Alto Genetics and Genomics Department
Roche Palo Alto
3431 Hillview Ave, Mailstop S3-1, Palo Alto, CA 94024, USA
Tel: 6508555807
Email: Jonathan.Usuka@roche.com
Primer A: No primer submitted with this STS
Primer B: No primer submitted with this STS.
FEATURES
source
1..627
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/map="1-21087-20591-CAAA01034157.1.1.28512"
/clone_lib="Roche Palo Alto"
/note="SNPs developed from assay sequences derived from 15
different strains of mice (as of October 1, 2003). Those
strains include A/J, A/HeJ, 129/Sv, AKR/J, B10.D2-H2/osnJ,
BALB/cByJ, BALB/cJ, C3H/HeJ, C57BL/6J, CAST/Ei, DBA/2J,
MRL/MpJ, NZB/BinJ, NZW/LacJ, SPRET/Ei."
<1..>627

STX
ORIGIN
Query Match      76.8%; Score 19.2; DB 10; Length 627;
Best Local Similarity 84.0%; Pred. No. 3.8e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAACACTTGAAGGGGAGACTCA 25
Db 306 ANACAACATCTGGAAGGGGAGACTCA 282

RESULT 35
RNTXALRC
LOCUS      689 bp mRNA linear ROD 18-APR-2005
DEFINITION Rat tissue mRNA for kallikrein C-terminal region (RC 3.4.21.8).
ACCESSION X03560
VERSION X03560.1 GI:57370
KEYWORDS kallikrein; protease; serine protease.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Rattus.
```



```
REFERENCE
AUTHORS      1 (bases 189 to 689)
TITLE        Gerald, W.L., Chao, J. and Chao, L.
JOURNAL      Immunological identification of rat tissue kallikrein cDNA and
PUBMED       characterization of the kallikrein gene family
PUBMED       Biochim. Biophys. Acta 866 (1), 1-14 (1986)
REFERENCE    2 (bases 1 to 188)
AUTHORS      Chao, L.
TITLE        Direct Submission
JOURNAL      Submitted (13-AUG-1986) Medical University of South Carolina, 171
JOURNAL      Ashley Avenue, Charleston, South Carolina 29425-2211
FEATURES
source       1..689
              Location/Qualifiers
              /organism="Rattus norvegicus"
              /mol_type="mRNA"
              /db_xref="taxon:10118"
              <1..646
              /codon_start=2
              /product="kallikrein"
              /protein_id="CAA27247.1"
              /db_xref="GI:818030"
              /db_xref="GOA:P00758"
              /db_xref="UniProt/Swiss-Prot:P00758"
              /translation="YLCGGVLIDPSWVITAAHCATDNYQVWLGRNNLYDEPPAQHRL
              VSQSFPHPGNQDLWNHTRPQGDYSNDMLHLSPADITDGVKVIDLPIEPPKVG
              STCLASGWSITPDGLESDLOCVNIDLLSNEKCVKAHKEVTDMLCAGEMDGGKD
              TCKGDSGGPLICNGVLQITSGWFGPCGEPKPGIYTKLIKFTPIKEVKNKP"
polyA_signal 670..675
              /note="putative"
polyA_site   689
ORIGIN
Query Match      76.8%; Score 19.2; DB 9; Length 689;
Best Local Similarity 87.5%; Pred. No. 3.8e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY      2 AAAAACTTGGGAAGGAGACTCA 25
      ||| ||||| ||||| ||||| |||||
Db      476 AAAGACACTTGCAGGGTGACTCA 499
      ||| ||||| ||||| ||||| |||||

RESULT 36
LOCUS      S48142      777 bp      mRNA      linear      ROD 08-MAY-1993
DEFINITION T-kininogenase=kallikrein homolog [rats, submandibular gland, mRNA,
ACCESSION  S48142
VERSION     S48142.1 GI:259430
KEYWORDS   Rattus sp.
SOURCE     Rattus sp.
ORGANISM   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
           Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE  1 (bases 1 to 777)
AUTHORS   Ma, J.X., Chao, J. and Chao, L.
TITLE     Molecular cloning and characterization of rKlk10, a cDNA encoding
JOURNAL   T-kininogenase from rat submandibular gland and kidney
PUBMED    Biochemistry 31 (44), 10922-10928 (1992)
REMARK    GenBank staff at the National Library of Medicine created this
entry [NCBI gisseq 118095] from the original journal article.
FEATURES
source     1..777
              Location/Qualifiers
              /organism="Rattus sp."
              /mol_type="mRNA"
              /db_xref="taxon:10118"
              1..777
              /gene="T-kininogenase"
              3..737
              /gene="T-kininogenase"
              /note="kallikrein homolog"
              /codon_start=1

Query Match      76.8%; Score 19.2; DB 9; Length 849;
Best Local Similarity 87.5%; Pred. No. 3.7e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY      2 AAAAACTTGGGAAGGAGACTCA 25
      ||| ||||| ||||| ||||| |||||
Db      567 AAAGACACTTGCAGGGTGACTCA 590
      ||| ||||| ||||| ||||| |||||

RESULT 37
LOCUS      MMNGFB      849 bp      mRNA      linear      ROD 30-MAR-1995
DEFINITION Mouse mRNA for gamma subunit of nerve growth factor (NGF).
ACCESSION  X01389
VERSION     X01389.1 GI:53373
KEYWORDS   Mus musculus (house mouse)
SOURCE     Mus musculus
ORGANISM   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
           Sciurognathi; Muridae; Murinae; Mus.
REFERENCE  1 (bases 1 to 849)
AUTHORS   Ullrich, A., Gray, A., Wood, W.I., Hayflick, J. and Seeburg, P.H.
TITLE     Isolation of a cDNA clone coding for the gamma-subunit of mouse
JOURNAL   nerve growth factor using a high-stringency selection procedure
PUBMED    DNA 3 (5), 387-392 (1984)
FEATURES
source     1..849
              Location/Qualifiers
              /organism="Mus musculus"
              /mol_type="mRNA"
              /db_xref="taxon:10090"
              18..803
              /note="unnamed protein product; gamma NGF precursor"
              /codon_start=1
              /protein_id="CAA25645.1"
              /db_xref="GI:53374"
              /db_xref="GOA:P00756"
              /db_xref="UniProt/Swiss-Prot:P00756"
              /translation="MMFLILFLALSLGGIDAAPPVQSVQRIVGGFKCKNSQPMHVVYR
              YTCGLCGVLLDPNMVLTAAHCYDDNYKVLGKNLFDKPSAQHRFVSKAIPHGPN
              MSIMRKHIRFLEVDYSDNLMRLSLKPADITDVKPITLPTBPKLQSTCLASGWSI
              TPTKFOFTDCLVCNKLILPNEDCAKAIKVTDMLCAGEMDGGKDTCKDGGSGPLI
              CDGLQITSGWHTPCGEPGIVTKLNKFTSWIKDTMAKNP"
              18..89
              /notes="put. signal peptide"
              90..800
              /product="gamma NGF"
              90..388
              /notes="proteolytic peptide B1"
              389..400
              /notes="connecting peptide"
              401..800
              /note="proteolytic peptide C2"
              825..830
              /notes="polyA signal"
              849
              /note="polyA site"
ORIGIN
Query Match      76.8%; Score 19.2; DB 9; Length 849;
Best Local Similarity 87.5%; Pred. No. 3.7e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY      2 AAAAACTTGGGAAGGAGACTCA 25
      ||| ||||| ||||| ||||| |||||
Db      567 AAAGACACTTGCAGGGTGACTCA 590
      ||| ||||| ||||| ||||| |||||

REFERENCE
AUTHORS      1 (bases 189 to 689)
TITLE        Gerald, W.L., Chao, J. and Chao, L.
JOURNAL      Immunological identification of rat tissue kallikrein cDNA and
PUBMED       characterization of the kallikrein gene family
PUBMED       Biochim. Biophys. Acta 866 (1), 1-14 (1986)
REFERENCE    2 (bases 1 to 188)
AUTHORS      Chao, L.
TITLE        Direct Submission
JOURNAL      Submitted (13-AUG-1986) Medical University of South Carolina, 171
JOURNAL      Ashley Avenue, Charleston, South Carolina 29425-2211
FEATURES
source       1..689
              Location/Qualifiers
              /organism="Rattus norvegicus"
              /mol_type="mRNA"
              /db_xref="taxon:10118"
              <1..646
              /codon_start=2
              /product="kallikrein"
              /protein_id="CAA27247.1"
              /db_xref="GI:818030"
              /db_xref="GOA:P00758"
              /db_xref="UniProt/Swiss-Prot:P00758"
              /translation="YLCGGVLIDPSWVITAAHCATDNYQVWLGRNNLYDEPPAQHRL
              VSQSFPHPGNQDLWNHTRPQGDYSNDMLHLSPADITDGVKVIDLPIEPPKVG
              STCLASGWSITPDGLESDLOCVNIDLLSNEKCVKAHKEVTDMLCAGEMDGGKD
              TCKGDSGGPLICNGVLQITSGWFGPCGEPKPGIYTKLIKFTPIKEVKNKP"
polyA_signal 670..675
              /note="putative"
polyA_site   689
ORIGIN
Query Match      76.8%; Score 19.2; DB 9; Length 689;
Best Local Similarity 87.5%; Pred. No. 3.8e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY      2 AAAAACTTGGGAAGGAGACTCA 25
      ||| ||||| ||||| ||||| |||||
Db      476 AAAGACACTTGCAGGGTGACTCA 499
      ||| ||||| ||||| ||||| |||||

RESULT 36
LOCUS      S48142      777 bp      mRNA      linear      ROD 08-MAY-1993
DEFINITION T-kininogenase=kallikrein homolog [rats, submandibular gland, mRNA,
ACCESSION  S48142
VERSION     S48142.1 GI:259430
KEYWORDS   Rattus sp.
SOURCE     Rattus sp.
ORGANISM   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
           Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE  1 (bases 1 to 777)
AUTHORS   Ma, J.X., Chao, J. and Chao, L.
TITLE     Molecular cloning and characterization of rKlk10, a cDNA encoding
JOURNAL   T-kininogenase from rat submandibular gland and kidney
PUBMED    Biochemistry 31 (44), 10922-10928 (1992)
REMARK    GenBank staff at the National Library of Medicine created this
entry [NCBI gisseq 118095] from the original journal article.
FEATURES
source     1..777
              Location/Qualifiers
              /organism="Rattus sp."
              /mol_type="mRNA"
              /db_xref="taxon:10118"
              1..777
              /gene="T-kininogenase"
              3..737
              /gene="T-kininogenase"
              /note="kallikrein homolog"
              /codon_start=1
```

```
QY 2 AAAAACTTGGAGGAGACTCA 25
||| ||||| ||||| |||||
Db 633 AAAGACACTTGCAGGGTGACTCA 656

RESULT 38
RATPKLK
DEFINITION rat pancreatic preprokallikrein mrna.
ACCESSION J00758
VERSION J00758.1 GI:206200
KEYWORDS kallikrein; protease; serine protease.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Murinae; Rattus.
1 (bases 1 to 867)
REFERENCE Swift,G.H., Degorn,J.C., Ashley,P.L., Cummings,S.W. and
AUTHORS MacDonald,R.J.
TITLE Rat pancreatic kallikrein mRNA: nucleotide sequence and amino acid
sequence of the encoded preproenzyme
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 79 (23), 7263-7267 (1982)
PUBMED 6961406
COMMENT Original source text: rat pancreatic tissue.
FEATURES
source
1..867
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"

ORIGIN
Query Match 76.8%; Score 19.2; DB 9; Length 867;
Best Local Similarity 87.5%; Pred. No. 3.6e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAACTTGGAGGAGACTCA 25
||| ||||| ||||| |||||
Db 652 AAAGACACTTGCAGGGTGACTCA 675

RESULT 39
RATKALPS
LOCUS Rat submaxillary gland PS kallikrein mRNA, complete cds.
DEFINITION Rat submaxillary gland PS kallikrein mRNA, complete cds.
ACCESSION M11563
VERSION M11563.1 GI:205029
KEYWORDS Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Murinae; Rattus.
1 (bases 1 to 873)
REFERENCE Ashley,P.L. and MacDonald,R.J.
AUTHORS Kallikrein-related mRNAs of the rat submaxillary gland: nucleotide
TITLE sequences of four distinct types including tonin
JOURNAL Biochemistry 24 (17), 4512-4520 (1985)
PUBMED 2998455
COMMENT Original source text: Rat submaxillary gland, cDNA to mRNA.
FEATURES
source
1..873
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
31..828
/notes="PS kallikrein"
/codon_start=1
/protein_id="AAA41464.1"
/db_xref="GI:205030"
/translacion="MPVTWFLILFLALSLGRNDAPPVQSRVVGYNCEMNSQPMQV
AVTFEYLCGGVLIDPSVITAAHCATONYQWLGRNNLYDEBPAPQRLVLSQSPFH

CDS
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E.B. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 184 Row: C Column: 17
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 13994130.
```

```
PGFNODLIWNHTROPDDYSNDLMLHLSPADITDGVKVIDLPIEKPQVSTCLASG
WGSITPDGLELSDLOQVNIIDLNLSEKCKVEAKKEVTDLMLCAGMDGCKTKGDSG
GPLICNGVLQIGITSWGFNPGCEPKPGIYTKLIKFTPIKVEVMKNP"

ORIGIN
Query Match 76.8%; Score 19.2; DB 9; Length 873;
Best Local Similarity 87.5%; Pred. No. 3.6e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAACTTGGAGGAGACTCA 25
||| ||||| ||||| |||||
Db 658 AAAGACACTTGCAGGGTGACTCA 681

RESULT 40
BC078784
LOCUS Rattus norvegicus nerve growth factor, gamma, mRNA (cDNA clone
DEFINITION MGC:93329 IMAGE:7128287), complete cds.
ACCESSION BC078784
VERSION BC078784.1 GI:51261179
KEYWORDS MGC.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Murinae; Rattus.
1 (bases 1 to 879)
REFERENCE Strausberg RL, Feingold EA, Grouse LH, Derge JG, Klausner RD,
AUTHORS Collins FS, Wagner L, Shenmen CM, Schuler GD, Altschul SF, Zeeberg
B, Buetow KH, Schaefer CF, Bhat NK, Hopkins RF, Jordan H, Moore T,
Max SI, Wang J, Hsieh F, Diatchenko L, Marusina K, Farmer AA, Rubin
GM, Hong L, Stapleton M, Soares MB, Donald MP, Casavant TL,
Schaeetz TE, Brownstein MJ, Udwin TB, Toshiyuki S, Carninci P,
Sj, Bosak SA, McKernan KJ, Malek JA, Gunaratne PH,
Richards S, Worley KC, Hale S, Garcia AM, Gay LJ, Hulyk SW,
Villalon DK, Muzny DM, Sodergren EJ, Lu X, Gibbs RA, Fahey J,
Heltan E, Kettelman M, Madan A, Rodrigues S, Sanchez A, Whiting M,
Madan A, Young AC, Shevchenko Y, Bouffard GG, Blakesley RW,
Touchman JW, Green ED, Dickinson MC, Rodriguez AC, Grimwood J,
Schmutz J, Myers RM, Butterfield YS, Krzywinski MI, Skalska U,
Smalusz DE, Schnerch A, Schein JE, Jones SJ and Marra MA.
Mammalian Gene Collection Program Team
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 879)
Director MGC Project.
Direct Submission
JOURNAL Submitted (02-AUG-2004) National Institutes of Health, Mammalian
PUBMED Gene Collection (MGC), Cancer Genomics Office, National Cancer
AUTHORS Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
JOURNAL USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgapbs@mail.nih.gov
Tissue Procurement: Howard Jacobs
cDNA Library Preparation: Express Genomics
cDNA Library Arrayed by: The I.M.A.G.E.B. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www-shgc.stanford.edu
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E.B. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 184 Row: C Column: 17
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 13994130.
```

FEATURES	Location/Qualifiers	CDS
source	1..879 /organism="Rattus norvegicus" /mol_type="mRNA" /db_xref="taxon:10116" /clone="MGC:93329 IMAGE:7128287" /tissue type="Kidney, rat (Brown Norway)" /clone_lib="NIH MGC_236" /lab_host="DH10E" /notes="Vector: pExpress1"	1..>1420 /gene="rbcl" /codon_start=1 /transl_table=11 /product="rubisco large subunit" /protein_id="CAA85682.1" /db_xref="GI:671609" /db_xref="GOA:Q36641" /db_xref="InterPro:IPR000685" /db_xref="UniProt/TREMBL:Q36641" /translation="MSPOTETKASVGFKAGVKEYKLTYYTPVETKDTDLAAFRVTP QVEPPEBAGAAVAESSTGTWTTVDGLTSLDRYKGRCYHIEPVPGEKQVLCVA YPLDFEGSVTNMFTSIVGNVGFPAKRALRLEDLRIIPAIKITQGEPHGLQVERD KLNKGRPLLGCTIKPKLGSARKYGRAVYECRLGLDFTKDDENVNSQPFMKWRDRF LFTCALYKQAQETGEIKGHYLNATAGTCEEMIKRFAFARLGLVPTVMHDYITGEFTA NTSLAHYCRDNGLLLIHRAHMAVIDROKNHGMHFRVLAKALRLSGNHIHSGTVVVK LEGEDITLGFVDLLRDDPLEKDRSGIVFTQDWYSLPSVIPVASGIIHVMMPALTE IFGDDSVLQFGGGTIGHPGNAPGAVANRVALEACVKAAREGRDLAREGNALIREACK WSPELAAACEVWKEIKFEPFAMD"
gene	1..879 /gene="Ngfg" /notes="synonyms: KUK1, Kk1c1, rGK-1, preprokallikrein" /db_xref="GeneID:24594" /db_xref="RATMAP:34215" /db_xref="RGD:3175"	
CDS	13..810 /gene="Ngfg" /codon_start=1 /product="nerve growth factor, gamma" /protein_id="AAH78784.1" /db_xref="GI:5161180" /db_xref="GeneID:24594" /db_xref="RATMAP:34215" /db_xref="RGD:3175" /translations="MPVTMMFLILFLALSLGRNDAAPPVQSRVVGVCNEMNSOPWQV AVYFGEYLGGVLIDPSVITAAHCATDNQVWLGNNLYDEPACQRLYQSOPFH PGFNQDLIWNHTKPGQDSDNMLHLSPADITDGVKVIDLPISPEPKVGSTCLASG WGSITPDGLESLDQCVNIDLLSNKCVKAEKBEVTDMLCAGEMDGGKOTCKGDSG GPLICNGVLQIGTSWGFNPGCEPKPGIYTKLIKFTPIKEVMKENP"	
ORIGIN		
Query Match	76.8%; Score 19.2; DB 9; Length 879;	Query Match 76.8%; Score 19.2; DB 15; Length 1420;
Best Local Similarity	87.5%; Pred. No. 3.6e+02;	Best Local Similarity 87.5%; Pred. No. 3.4e+02;
Matches	21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;	Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy	2 AAAAAACACTTGGGAAGGAGACTCA 25 	2 AAAAAACACTTGGGAAGGAGACTCA 25
Db	640 AAGACACTTGCAGGCTGACTCA 663 	1125 AATACACTTGCAGGAGACCCA 1102
RESULT 41		RESULT 42
PHRCL52/c		AB036705
LOCUS	PHRCL52 1420 bp DNA linear PLN 18-APR-2005	AB036705 46235 bp DNA linear PRI 06-JAN-2004
DEFINITION	P.hysopifolia chloroplast rbcl gene for rubisco (large subunit) (partial).	Homo sapiens EMS1 gene for cortactin, partial cds.
ACCESSION	Z37432	AB036705
VERSION	237432.1 GI:671608	AB036705.1 GI:40645043
KEYWORDS	rbcl gene; Rubisco.	
SOURCE	chloroplast Prunella hyssopifolia	Homo sapiens (human)
ORGANISM	Prunella hyssopifolia	Homo sapiens
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; asterids; lamids; Lamiales; Lamiaceae; Nepetoideae; Nepetaceae; Prunella.	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 1420) Kaufmann,M. and Wink,M. Molecular systematics of the Nepetoideae (family Labiatae): phylogenetic implications from rbcl gene sequences Biosci. Rep. 49, 635-645 (1994)	1 Yamashita,A. and Togashi,S. Human EMS1(cortactin) gene Published only in Database (2004) 2 (bases 1 to 46235) Yamashita,A. and Togashi,S. Direct Submission Submitted (06-JAN-2000) Shin Togashi, Kitasato University, School of Science, Dept. of Biosciences, Kitasato 1-15-1, Sagamihara, Kanagawa 228-8555, Japan (E-mail:stogashi@jet.sci.kitasato-u.ac.jp, Tel:81-42-778-9481, Fax:81-42-778-9481) Location/Qualifiers 1..46235 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" Join(15820..15906,16040..16113,18353..18482,23064..23174,24185..24239,25535..25645,28268..28378,28922..29032,29993..30103,31463..31518) /gene="EMS1" Join(15820..15906,16040..16113,18353..18482,23064..23174,24185..24239,25535..25645,28268..28378,28922..29032,29993..30103,31463..31518) /gene="EMS1" /codon_start=1 /product="cortactin" /protein_id="BAD06416.1" /db_xref="GI:40645044"
FEATURES	Location/Qualifiers	FEATURES
source	1..1420 /organism="Prunella hyssopifolia" /organelle="plastid:chloroplast" /mol_type="genomic DNA" /db_xref="taxon:39357" 1..>1420 /gene="rbcl"	source 1..46235 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" Join(15820..15906,16040..16113,18353..18482,23064..23174,24185..24239,25535..25645,28268..28378,28922..29032,29993..30103,31463..31518) /gene="EMS1" Join(15820..15906,16040..16113,18353..18482,23064..23174,24185..24239,25535..25645,28268..28378,28922..29032,29993..30103,31463..31518) /gene="EMS1" /codon_start=1 /product="cortactin" /protein_id="BAD06416.1" /db_xref="GI:40645044" QSGSHOEINIHKLRNVEHQHLETKGKSHGYGKFGVGEQDRMDKSAVGH EYQSKLSKHCQSQVDSVRGFGKGQMDRVQSAVGFEYQKTEKIASQKDYSSGFGG
gene	1..>1420 /gene="rbcl"	

KYGVQADRVDSAVGQYQKTEKHESORDYSKFGGKYGIDKDKVDKSAVGFYQOGK
TEKHESQKDYKVGKFGKGVQDRODKALGDHDEKQLHESQKDYKTFGGKRGVQ
SERQDSAAVGFDPYKEKLAKHESQQDYSKFGGKYGKQKDRMDK*

ORIGIN

Query Match 76.8%; Score 19.2; DB 8; Length 46235;
Best Local Similarity 87.5%; Pred. No. 2.1e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 AAAAACTTGAAGGGAGACTCA 25
|||||
Db 42251 AAAAGCACTTGAAGGGAGACCCA 42274
|||||

RESULT 43

AC101212
LOCUS Mus musculus clone RP23-182G21, LOW-PASS SEQUENCE SAMPLING.
AC101212
ACCESSION
VERSION AC101212.1 GI:17059987
KEYWORDS HTG; HTGS, PHASE0.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Murioidea; Muridae; Murinae; Mus.
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
1 (bases 1 to 59388)
Mus musculus, clone RP23-182G21
Unpublished
2 (bases 1 to 59388)

REFERENCE

AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barina,N., Bastien,V., Boguslavsky,I., Boukhgalter,B.,
Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
Chospel,Y., Colangelo,M., Collins,S., Collamore,A., Cook,A.,
Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K.,
Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G.,
McClean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J.,
Meneus,L., Mihova,T., Mienga,V., Murphy,T., Naylor,J., Nguyen,C.,
Nebus,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

TITLE

JOURNAL
COMMENT
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L16243
Center Clone name: 182_G_21

* NOTE: This record contains 76 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for

* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

1 677: contig of 677 bp in length
678 777: gap of 100 bp
778 1490: contig of 713 bp in length
778 1490: contig of 713 bp in length
1491 1590: gap of 100 bp
1591 2295: contig of 705 bp in length
2296 2395: gap of 100 bp
2396 3101: contig of 706 bp in length
3102 3201: gap of 100 bp
3202 3880: contig of 679 bp in length
3881 3980: gap of 100 bp
3981 4643: contig of 663 bp in length
4644 4743: gap of 100 bp
4744 5399: contig of 656 bp in length
5400 5499: gap of 100 bp
5500 6174: contig of 675 bp in length
6175 6274: gap of 100 bp
6275 6966: contig of 692 bp in length
6967 7066: gap of 100 bp
7067 7732: contig of 666 bp in length
7733 7832: gap of 100 bp
7833 8537: contig of 705 bp in length
8538 8637: gap of 100 bp
8638 9326: contig of 689 bp in length
9327 9426: gap of 100 bp
9427 10125: contig of 699 bp in length
10126 10225: gap of 100 bp
10226 10894: contig of 669 bp in length
10895 10994: gap of 100 bp
10995 11657: contig of 663 bp in length
11658 11757: gap of 100 bp
11758 12423: contig of 666 bp in length
12424 12523: gap of 100 bp
12524 13209: contig of 686 bp in length
13210 13309: gap of 100 bp
13310 13997: contig of 688 bp in length
13998 14097: gap of 100 bp
14098 14786: contig of 689 bp in length
14787 14886: gap of 100 bp
14887 15880: contig of 694 bp in length
15881 16374: contig of 694 bp in length
16375 16474: gap of 100 bp
16475 17176: contig of 702 bp in length
17177 17276: gap of 100 bp
17277 17943: contig of 667 bp in length
17944 18043: gap of 100 bp
18044 18666: contig of 623 bp in length
18667 18766: gap of 100 bp
18767 19427: contig of 661 bp in length
19428 19527: gap of 100 bp
19528 20201: contig of 674 bp in length
20202 20301: gap of 100 bp
20302 20961: contig of 660 bp in length
20962 21061: gap of 100 bp
21062 21753: contig of 692 bp in length
21754 21853: gap of 100 bp
21854 22550: contig of 697 bp in length
22551 22650: gap of 100 bp
22651 23301: contig of 651 bp in length
23302 23401: gap of 100 bp
23402 24085: contig of 684 bp in length
24086 24185: gap of 100 bp
24186 24894: contig of 709 bp in length
24895 24994: gap of 100 bp
24995 25691: contig of 697 bp in length
25692 25791: gap of 100 bp
25792 26465: contig of 674 bp in length

```

* 26466 26565: gap of 100 bp
* 26566 contig of 691 bp in length
* 27257 27356: gap of 100 bp
* 27357 contig of 702 bp in length
* 28059 28158: gap of 100 bp
* 28159 contig of 679 bp in length
* 28838 28937: gap of 100 bp
* 28938 contig of 709 bp in length
* 29647 29746: gap of 100 bp
* 29747 contig of 689 bp in length
* 30436 30535: gap of 100 bp
* 30536 contig of 662 bp in length
* 31198 31297: gap of 100 bp
* 31298 contig of 663 bp in length
* 31961 32060: gap of 100 bp
* 32061 contig of 664 bp in length
* 32725 32824: gap of 100 bp
* 32825 contig of 672 bp in length
* 33497 33596: gap of 100 bp
* 33597 contig of 689 bp in length
* 34286 34385: gap of 100 bp
* 34386 contig of 683 bp in length
* 35069 35168: gap of 100 bp
* 35169 contig of 679 bp in length
* 35948 35947: gap of 100 bp
* 35948 contig of 685 bp in length
* 36633 36732: gap of 100 bp
* 36733 contig of 665 bp in length
* 37398 37497: gap of 100 bp
* 37498 contig of 696 bp in length
* 38194 38293: gap of 100 bp
* 38294 contig of 717 bp in length
* 39011 39110: gap of 100 bp
* 39111 contig of 687 bp in length
* 39798 39897: gap of 100 bp
* 39898 contig of 663 bp in length
* 40561 40660: gap of 100 bp
* 40661 contig of 687 bp in length
* 41348 41447: gap of 100 bp
* 41448 contig of 685 bp in length
* 42133 42232: gap of 100 bp
* 42233 contig of 685 bp in length
* 42918 43017: gap of 100 bp
* 43018 contig of 701 bp in length
* 43719 43818: gap of 100 bp
* 43819 contig of 695 bp in length
* 44513 44613: gap of 100 bp
* 44613 contig of 680 bp in length
* 45293 45393: gap of 100 bp
* 45394 contig of 726 bp in length
* 46120 46219: gap of 100 bp
* 46220 contig of 649 bp in length
* 46869 46968: gap of 100 bp
* 46969 contig of 657 bp in length
* 47626 47725: gap of 100 bp
* 47726 contig of 664 bp in length
* 48330 48430: gap of 100 bp
* 48430 contig of 681 bp in length
* 49171 49270: gap of 100 bp
* 49270 contig of 698 bp in length
* 49969 50068: gap of 100 bp
* 50069 contig of 681 bp in length
* 50750 50849: gap of 100 bp
* 50850 contig of 712 bp in length
* 51562 51661: gap of 100 bp
* 51662 contig of 650 bp in length
* 52312 52411: gap of 100 bp
* 52412 contig of 705 bp in length
* 53117 53216: gap of 100 bp
* 53217 contig of 718 bp in length

```

Query Match 76.8%; Score 19.2; DB 14; Length 59388;
 Best Local Similarity 87.5%; Pred. No. 2e+02;

```

Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 2 AAAAACACTTGGAGGGAGACTCA 25
    |||||
Db 13725 AAAAACACTTGGACACAGACTGA 13748
    |||||

RESULT 44
AC123415.2
WPCOMMENT
Sequence split into 5 fragments LOCUS AC123415 Accession AC123415
Fragment Name Begin End
AC123415_0 1 110000
AC123415_1 100001 210000
AC123415_2 200001 310000
AC123415_3 300001 410000
AC123415_4 400001 421216
Continuation (3 of 5) of AC123415 from base 200001 (AC123415 Rattus norvegicus clone CH

Query Match 76.8%; Score 19.2; DB 14; Length 110000;
Best Local Similarity 87.5%; Pred. No. 1.9e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 2 AAAAACACTTGGAGGGAGACTCA 25
    |||||
Db 50734 AAAAACACTTGGAAAGTGATTCA 50757
    |||||

RESULT 45
AC159696
LOCUS
DEFINITION
    AC159696 125379 bp DNA linear HTG 01-JUL-2005
    Bos taurus clone CH240-75G13, WORKING DRAFT SEQUENCE, 9 unordered
    pieces.
ACCESSION
    AC159696
VERSION
    AC159696.3 GI:68299868
KEYWORDS
    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE
    Bos taurus (cow)
ORGANISM
    Bos taurus
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
    Pecora; Bovidae; Bovinae; Bos.
    Muzny,D.Marie., Metzker,M.Lee., Abramson,S., Adams,C., Alder,J.,
    Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
    Anyalebechi,V., Ayodeji,A., Ayodeji,M., Baca,E., Baden,H.,
    Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
    Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
    Bryant,N., Buhay,C., Burch,P., Burrell,K., Ceasar,H., Center,A.,
    Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Chen,Y., Chen,Z.,
    Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
    Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
    Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
    Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinu,H., Divya,K.,
    Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Evans,K.,
    Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
    Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
    Fraser,C.M., Gabisi,A., Gantä,R., Garcia,A., Garner,T., Garza,M.,
    Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,M., Guevara,W.,
    Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., HamilCon,K.,
    Hernandez,R., Hines,S., Hladun,S.L., Hodgson,N., Hernandez,J.,
    Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
    Jarkpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
    Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
    Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
    Lorensuhewa,L., Loulseghe,H., Lozado,R.J., Lu,X., Ma,J.,
    Maheshwari,M., Mahindartine,M., Mahmoud,M., Malloy,K., Mangum,A.,
    Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
    Mawhiney,S., Mcleod,M.P., McNeill,T.Z., Meenen,E.,
    Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
    Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
    Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,

```


Japan (E-mail: hattori@sc.riken.go.jp,
URL: http://hgp.qsc.riken.go.jp/, Tel: 81-42-778-9923,
Fax: 81-42-778-9924)
On May 31, 2000 this sequence version replaced gi:6997436.

COMMENT

----- Genome Center
Center: RIKEN Genomic Sciences Center (GSC)
Center code: RIKEN
Web site: http://hgp.qsc.riken.go.jp/
Contact: hattori@sc.riken.go.jp
----- Project Information
Center project name: Humdraft11
Center clone name: CMB9-16D10
----- Summary Statistics
Sequencing vector: PCR products; 100% of reads
Chemistry: Dye-terminator ET-amersham; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 122741 bases at least Q40
Consensus quality: 131182 bases at least Q30
Consensus quality: 135698 bases at least Q20
Insert size: 138675; sum-of-contigs
Quality coverage: 4.07x in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of 31 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

```
1 24550 contig of 24550 bp in length
24651 34442 contig of 9792 bp in length
34543 43950 contig of 9408 bp in length
44051 51303 contig of 7253 bp in length
51404 58548 contig of 7145 bp in length
58649 64379 contig of 5731 bp in length
64480 71839 contig of 7360 bp in length
71940 76893 contig of 4853 bp in length
76893 82501 contig of 5609 bp in length
82602 86341 contig of 3740 bp in length
86442 91700 contig of 5259 bp in length
91801 97119 contig of 5319 bp in length
97220 100996 contig of 3777 bp in length
101097 106357 contig of 5261 bp in length
106458 110874 contig of 4417 bp in length
110975 114272 contig of 3453 bp in length
114528 118179 contig of 3652 bp in length
118280 121483 contig of 3203 bp in length
121583 124342 contig of 2760 bp in length
124443 126512 contig of 2070 bp in length
126513 129400 contig of 2788 bp in length
129401 132503 contig of 1665 bp in length
132504 133166 contig of 100 bp in length
133166 132521 contig of 1255 bp in length
132521 132621 contig of 100 bp in length
132621 133646 contig of 1026 bp in length
133647 133746 contig of 1157 bp in length
133747 134903 contig of 1039 bp in length
134904 136042 contig of 1039 bp in length
136043 137315 contig of 1173 bp in length
137316 137415 contig of 1194 bp in length
137416 138609 contig of 100 bp in length
138610 138709 contig of 1183 bp in length
138710 139892 contig of 100 bp in length
139893 141097 contig of 1105 bp in length
141098 141197 contig of 100 bp in length
141198 141675 contig of 478 bp in length
```

Sequence updated (12-Oct-1999)

Sequence updated (26-May-2000).

* NOTE: This is a 'working draft' sequence. It currently consists of 31 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

```
1 24550: contig of 24550 bp in length
24551 24650: gap of 100 bp
24651 34442: contig of 9792 bp in length
```

```
* 34443 34542: gap of 100 bp
34543 43950: contig of 9408 bp in length
43951 44050: gap of 100 bp
44051 51303: contig of 7253 bp in length
51304 51403: gap of 100 bp
51404 58548: contig of 7145 bp in length
58549 58649: gap of 100 bp
58649 64379: contig of 5731 bp in length
64380 64479: gap of 100 bp
64480 71839: contig of 7360 bp in length
71840 71939: gap of 100 bp
71940 76792: contig of 4853 bp in length
76793 76892: gap of 100 bp
76893 82501: contig of 5609 bp in length
82502 82601: gap of 100 bp
82602 86341: contig of 3740 bp in length
86342 86441: gap of 100 bp
86442 91700: contig of 5259 bp in length
91701 91800: gap of 100 bp
91801 97119: contig of 5319 bp in length
97120 97219: gap of 100 bp
97220 100996: contig of 3777 bp in length
100997 101096: gap of 100 bp
101097 106357: contig of 5261 bp in length
106358 106457: gap of 100 bp
106458 110874: contig of 4417 bp in length
110875 110974: gap of 100 bp
110975 114272: contig of 3453 bp in length
114273 114527: gap of 100 bp
114528 118179: contig of 3652 bp in length
118180 118279: gap of 100 bp
118280 121482: contig of 3203 bp in length
121483 121582: gap of 100 bp
121583 124342: contig of 2760 bp in length
124343 124442: gap of 100 bp
124443 126512: contig of 2070 bp in length
126513 129400: gap of 100 bp
129401 129500: contig of 2788 bp in length
129501 131185: contig of 1665 bp in length
131186 131285: gap of 100 bp
131286 132520: contig of 1255 bp in length
132521 132620: gap of 100 bp
132621 133646: contig of 1026 bp in length
133647 133746: gap of 100 bp
133747 134903: contig of 1157 bp in length
134904 135003: gap of 100 bp
135004 136042: contig of 1039 bp in length
136043 137315: contig of 1173 bp in length
137316 137415: gap of 100 bp
137416 138609: contig of 1194 bp in length
138610 138709: gap of 100 bp
138710 139892: contig of 1183 bp in length
139893 139992: gap of 100 bp
139993 141097: contig of 1105 bp in length
141098 141197: gap of 100 bp
141198 141675: contig of 478 bp in length.
```

FEATURES

```
source
1. .141675
   /organism="Homo sapiens"
   /mol_type="genomic DNA"
   /db_xref="taxon:9606"
   /chromosome="11"
   /map="11q13"
   /clone="CMB9-16D10"
misc_feature
1. .24550
   /note="assembly_fragment"
misc_feature
24651. .34442
   /note="assembly_fragment"
misc_feature
34543. .43950
   /note="assembly_fragment"
misc_feature
44051. .51303
```

```
misc_feature /note="assembly_fragment"
51404..58548
/note="assembly_fragment clone_end:T7 vector_side:right"
misc_feature 58649..64379
/note="assembly_fragment"
misc_feature 64480..71839
/note="assembly_fragment"
misc_feature 71940..76792
/note="assembly_fragment"
misc_feature 76893..82501
/note="assembly_fragment"
misc_feature 82602..86341
/note="assembly_fragment"
misc_feature 86442..91700
/note="assembly_fragment"
misc_feature 91801..97119
/note="assembly_fragment"
misc_feature 97220..100996
/note="assembly_fragment"
misc_feature 101097..106357
/note="assembly_fragment"
misc_feature 106458..110874
/note="assembly_fragment"
misc_feature 110975..114427
/note="assembly_fragment"
misc_feature 114528..118179
/note="assembly_fragment"
misc_feature 118280..121482
/note="assembly_fragment"
misc_feature 121583..124342
/note="assembly_fragment"
```

```
Query Match 76.8%; Score 19.2; DB 14; Length 141675;
Best Local Similarity 87.5%; Pred. No. 1.8e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAAACACTTGGAGGGAGACTCA 25
Db 101734 AAAGCATTGGAGGGAGACCCA 101711
|||||
|||||
|||||
|||||
|||||
```

```
RESULT 47
AC073631 155877 bp DNA linear HTG 27-JUN-2000
LOCUS Homo sapiens chromosome 3 clone RP11-195H14, *** SEQUENCING IN
DEFINITION PROGRESS ***, 13 unordered pieces.
ACCESSION AC073631.1 GI:8748850
VERSION HTG; HTGS_PHASE1.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 155877)
AUTHORS Waterston,R.H.
TITLE The sequence of Homo sapiens clone
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 155877)
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (27-JUN-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
```

```
COMMENT
----- Genome Center -----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 13 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
```

```
* be preserved.
* 1 1086: contig of 1086 bp in length
* 1087 1186: gap of unknown length
* 1187 2446: contig of 1260 bp in length
* 2447 2546: gap of unknown length
* 2547 5563: contig of 3017 bp in length
* 5564 5663: gap of unknown length
* 5664 9561: contig of 3898 bp in length
* 9562 9661: gap of unknown length
* 9662 18553: contig of 8892 bp in length
* 18554 18653: gap of unknown length
* 18654 29773: contig of 11120 bp in length
* 29774 29873: gap of unknown length
* 29874 42605: contig of 12732 bp in length
* 42606 42706: gap of unknown length
* 42707 56734: contig of 14029 bp in length
* 56735 56835: gap of unknown length
* 56836 70735: contig of 13900 bp in length
* 70736 70835: gap of unknown length
* 70836 85357: contig of 14523 bp in length
* 85358 85458: gap of unknown length
* 85459 101145: contig of 15688 bp in length
* 101146 101246: gap of unknown length
* 101247 121898: contig of 20653 bp in length
* 121899 121999: gap of unknown length
* 122000 155877: contig of 33879 bp in length.
Location/Qualifiers
1..155877
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="3"
/clone="RP11-195H14"
1087..1186
/estimated_length=unknown
2447..2546
/estimated_length=unknown
5564..5663
/estimated_length=unknown
9562..9661
/estimated_length=unknown
18554..18653
/estimated_length=unknown
29774..29873
/estimated_length=unknown
42606..42706
/estimated_length=unknown
56735..56835
/estimated_length=unknown
70736..70835
/estimated_length=unknown
85358..85458
/estimated_length=unknown
101146..101246
/estimated_length=unknown
121899..121999
/estimated_length=unknown
```

ORIGIN

```
Query Match 76.8%; Score 19.2; DB 14; Length 155877;
Best Local Similarity 87.5%; Pred. No. 1.8e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAACACTTGGAGGGAGACTC 24
Db 64740 AGAAAAACACTTGGAGGGAGACTC 64763
|||||
|||||
|||||
|||||
|||||

RESULT 48
AC161483 161497 bp DNA linear HTG 15-MAY-2005
LOCUS Mus musculus chromosome 5 clone RP23-469M13 map 5, WORKING DRAFT
DEFINITION SEQUENCE, 21 unordered pieces.
```


gap 9983..110082
/estimated_length=unknown
misc_feature 10083..11832
/note="assembly_fragment"
gap 11833..11932
/estimated_length=unknown
misc_feature 11933..14903
/note="assembly_fragment"
gap 14904..15003
/estimated_length=unknown
misc_feature 15004..18537
/note="assembly_fragment"
gap 18538..18637
/estimated_length=unknown
misc_feature 18638..21921
/note="assembly_fragment"
gap 21922..22021
/estimated_length=unknown
misc_feature 22022..26145
/note="assembly_fragment"
gap 26146..26245
/estimated_length=unknown
misc_feature 26246..31488
/note="assembly_fragment"
gap 31489..31588
/estimated_length=unknown
misc_feature 31589..33657
/note="assembly_fragment"
gap 33658..36667
/estimated_length=unknown
misc_feature 36668..42840
/note="assembly_fragment"
gap 42841..42940
/estimated_length=unknown
misc_feature 42941..70982
/note="assembly_fragment"
gap 70983..71082
/estimated_length=unknown
misc_feature 71083..87513
/note="assembly_fragment"
gap 87514..87613
/estimated_length=unknown
misc_feature 87614..104740
/note="assembly_fragment"
gap 104741..104840
/estimated_length=unknown
misc_feature 104841..128106
/note="assembly_fragment"
gap 128107..128206
/estimated_length=unknown
misc_feature 128207..154032

Query Match 76.8%; Score 19.2; DB 14; Length 161497;
Best Local Similarity 87.5%; Pred. No. 1.8e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAACACTTGGAGGAGACTC 24
||||| ||||| ||||| ||||| |||||
Db 78454 AAAAAACACTGGGAGTGATACCTC 78477

RESULT 49
AC114655/c
LOCUS AC114655 162542 bp DNA linear ROD 04-SRP-2004
DEFINITION Mus musculus chromosome 1, clone RP24-160J12, complete sequence.
ACCESSION AC114655
VERSION AC114655.12 GI:51889857
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS

1 (bases 1 to 162542)
Birren,B., Nusbaum,C. and Lander,E.
Mus musculus chromosome 1, clone RP24-160J12
Unpublished
2 (bases 1 to 162542)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L.,
Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,
Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
Landers,T., Lehoczy,J., Levine,R., Lindblad-Toh,K., Liu,G.,
MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L.,
Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE
JOURNAL
REFERENCE
AUTHORS

Direct Submission
Submitted (10-MAR-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 162542)
Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V.,
Bloom,T., Boguslavskiy,L., Boukhgalter,B., Camarata,J., Chang,J.,
Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B.,
DeArellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L.,
Erickson,J., Faro,S., Ferreira,P., FitzGerald,M., Gage,D.,
Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N.,
Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I.,
Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T.,
Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbitt,R.,
MacLean,C., Macdonald,P., Major,J., Manning,J., Matthews,C.,
McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
Murphy,T., Naylor,J., Nguyen,C., Nguyen,T., Nicol,R., Norbu,C.,
O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C.,
Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Vassiliev,H., Venkataranan,V.S., Viel,R.,
Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L.,
Zimmer,A. and Zody,M.

TITLE
JOURNAL
REFERENCE
AUTHORS

Direct Submission
Submitted (03-AUG-2004) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 162542)
Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V.,
Bloom,T., Boguslavskiy,L., Boukhgalter,B., Camarata,J., Chang,J.,
Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B.,
DeArellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L.,
Erickson,J., Faro,S., Ferreira,P., FitzGerald,M., Gage,D.,
Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N.,
Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I.,
Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T.,
Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbitt,R.,
MacLean,C., Macdonald,P., Major,J., Manning,J., Matthews,C.,
McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
Murphy,T., Naylor,J., Nguyen,C., Nguyen,T., Nicol,R., Norbu,C.,
O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C.,
Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,

Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J.,
Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R.,
Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L.,
Zimmer, A. and Zody, M.

TITLE

Direct Submission
Submitted (04-SEP-2004) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA

COMMENT

On Sep 4, 2004 this sequence version replaced gi:50897427.
All repeats were identified using RepeatMasker:
Smith, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/MIT Center for Genome Research

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence_submissions@broad.mit.edu

----- Project Information

Center project name: L24414

Center clone name: 160_J_12

FEATURES

source

Location/Qualifiers

1..162542
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="1"
/map="1"
/clone="RP24-160J12"
/clone_lib="RPCI-24 Male Mouse BAC"

misc_feature

1..4
/notes="clone boundary
clone end,SP6
site:MboI"

repeat_region

complement(2701..2910)

/rpt_family="Lx3"

complement(2912..3061)

/rpt_family="L1M1"

3243..3832

/rpt_family="Lx9"

3833..4198

/rpt_family="RMER19B"

4404..4472

/rpt_family="T-rich"

complement(4486..4610)

/rpt_family="B1F"

complement(5657..5811)

/rpt_family="RSINE1"

5991..6234

/rpt_family="ORR1B1"

complement(6237..6315)

/rpt_family="L1_MM"

6313..6557

/rpt_family="L1_MM"

6637..6692

/rpt_family="ORR1B1"

complement(6749..6870)

/rpt_family="B1F"

7031..7231

/rpt_family="MT2B"

complement(7225..7284)

/rpt_family="ORR1D"

7285..7346

/rpt_family="TD_B1"

7298..7356

/rpt_family="PB1D7"

complement(7440..7651)

/rpt_family="B3A"

8046..8188

/rpt_family="B1_MM"

8191..8259

/rpt_family="(A)n"

complement(8751..9118)

/rpt_family="MTC"

repeat_region 9927..10057
/rpt_family="B1F"
repeat_region 10092..10187
/rpt_family="B1F"
repeat_region complement(10341..10493)
/rpt_family="B1_MM"
repeat_region 10714..10938
/rpt_family="B3"
repeat_region 11020..11114
/rpt_family="PB1D7"
repeat_region complement(11115..11428)
/rpt_family="MTB"
repeat_region complement(11780..11988)
/rpt_family="RMER4B"
repeat_region 11989..12018
/rpt_family="(TTA)n"
repeat_region complement(12019..12491)
/rpt_family="Lx2"
repeat_region complement(12493..12628)
/rpt_family="RMER4B"
repeat_region complement(12723..13006)
/rpt_family="RMER4B"
repeat_region complement(13033..13086)
/rpt_family="PB1D9"
repeat_region 13101..13238
/rpt_family="CT-rich"
repeat_region complement(13239..13371)
/rpt_family="B1F"
repeat_region complement(13377..13451)
/rpt_family="PB1D7"
repeat_region complement(13571..15097)
/rpt_family="L1_MM"
repeat_region complement(15213..15717)
/rpt_family="L1"
repeat_region complement(15727..18178)
/rpt_family="LLME"
repeat_region 18179..18213
/rpt_family="(TG)n"
repeat_region complement(18214..18820)
/rpt_family="LLME"
repeat_region 20168..20317
/rpt_family="B1F"
repeat_region complement(20523..20672)
/rpt_family="MIR"

Query Match 76.8%; Score 19.2; DB 9; Length 162542;

Best Local Similarity 87.5%; Pred.No.1.8e+02;

Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAAACACTTGGAGGAGACTCA 25

Db 120881 AAAAAACACTTGGAGGAGACTCA 120858

RESULT 50

AL590406

DEFINITION Human DNA sequence from clone Rp11-418P12 on chromosome 6, complete

sequence.

ACCESSION AL590406

VERSION AL590406.6 GI:14422236

KEYWORDS HTG; HMGCL;

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Metheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

1 (bases 1 to 164831)

Smith, M.

Direct Submission

Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,

Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vega@sanger.ac.uk

Clone requests: clonerequest@sanger.ac.uk

COMMENT

On Jun 13, 2001 this sequence version replaced gi:14270001.
 The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:
 Em:, EMBL; Sw:, SWISSPROT; Tr:, TrEMBL; Wp:, WormPEP; Information on the WormPEP database can be found at
http://www.sanger.ac.uk/projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping Group. Further information can be found at
<http://www.sanger.ac.uk/HGP/Chr6>
 RP11-418P12 is from the library RPCI-11.2 constructed by the group of Peter de Jong. For further details see
<http://www.chori.org/bacpac/home.htm>
 VECTOR: pBACE3.6

----- Genome Center
 Center: Wellcome Trust Sanger Institute
 Center code: SC
 Web site: <http://www.sanger.ac.uk>
 Contact: vega@sanger.ac.uk

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

FEATURES

source

1. .164831
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="R2PD:RPCIB753P12418"
 /db_xref="taxon:9606"
 /chromosome="6"
 /clone="RP11-418P12"
 /clone_lib="RPCI-11.2"

misc_feature

1
 /note="Clone_left_end: RP11-418P12"
 join(complement(AL590290.5:2101..2252),
 complement(151703..151800),complement(129944..130024),
 complement(129613..129720),complement(104392..104487),
 complement(101932..102080),complement(87120..87183),
 complement(83303..83491),complement(27318..27443),
 complement(22254..23647))
 /locus_tag="RP11-418P12.1-002"
 join(complement(AL590290.5:2101..2252),
 complement(151703..151800),complement(129944..130024),
 complement(129613..129720),complement(104392..104487),
 complement(101932..102080),complement(87120..87183),
 complement(83303..83491),complement(27318..27443),
 complement(22254..23647))
 /locus_tag="RP11-418P12.1-002"
 /note="match: cDNAs: AK055075.1"

mRNA

1
 join(complement(AL590290.5:2101..2247),
 complement(129944..130024),complement(129613..129720),
 complement(104392..104487),complement(101932..102080),
 complement(87120..87183),complement(83303..83491),
 complement(27318..27443),complement(22254..23647))
 /locus_tag="RP11-418P12.1-001"
 join(complement(AL590290.5:2101..2247),
 complement(129944..130024),complement(129613..129720),
 complement(104392..104487),complement(101932..102080),
 complement(87120..87183),complement(83303..83491),
 complement(27318..27443),complement(22254..23647))
 /locus_tag="RP11-418P12.1-001"

gene

1
 join(complement(AL590290.5:2101..2247),
 complement(129944..130024),complement(129613..129720),
 complement(104392..104487),complement(101932..102080),
 complement(87120..87183),complement(83303..83491),
 complement(27318..27443),complement(22254..23647))
 /locus_tag="RP11-418P12.1-001"
 join(complement(AL590290.5:2101..2247),
 complement(129944..130024),complement(129613..129720),
 complement(104392..104487),complement(101932..102080),
 complement(87120..87183),complement(83303..83491),
 complement(27318..27443),complement(22254..23647))
 /locus_tag="RP11-418P12.1-001"

mRNA

1
 join(complement(AL590290.5:2101..2247),
 complement(129944..130024),complement(129613..129720),
 complement(104392..104487),complement(101932..102080),
 complement(87120..87183),complement(83303..83491),
 complement(27318..27443),complement(22254..23647))
 /locus_tag="RP11-418P12.1-001"
 /note="match: ESTs: BG664787 BG703604 BI824058 F06678
 match: cDNAs: AF131827.1 BI824058"

polyA_site

1
 /locus_tag="RP11-418P12.1-001"

polyA_signal

1
 /locus_tag="RP11-418P12.1-001"

CDS

1
 complement(join(23546..23647,27318..27443,83303..83491,
 87120..87183,101932..102080,104392..104487))

/locus_tag="RP11-418P12.1-001"
 /standard_name="OTTHUMP0000016648"
 /codon_start=1
 /protein_id="CAI40662.1"
 /db_xref="GI:57209295"
 /db_xref="GOA:Q5JRJ4"
 /db_xref="InterPro:IPR000891"
 /db_xref="UniProt/TrEMBL:Q5JRJ4"
 /translation="MADTEVMKGIHOYPGVRYPLTPNLQGFHHAAGAATISVFG
 AASFSKKNINCISIESMGKPEEVVKSARHNNIPARGYVSCALGCPVEGSIIPKQVT
 EVSRRLXGMGCYEISLGTIGVTPGSMKRMLESVMKGIFFGALAVHCHDYTGQALAN
 ILTALQMGINVDYSAVSGLGCPYAKGASGNVATEDLLIYMLNGLGINTGVNLYKVEA
 GDFICKAVNTTKSKVAQAFNA"
 join(complement(AL590290.5:2104..2226),
 complement(129944..130027),complement(129613..129720),
 complement(99959..102080))
 /locus_tag="RP11-418P12.1-003"
 join(complement(AL590290.5:2104..2226),
 complement(129944..130027),complement(129613..129720),
 complement(99959..102080))
 /locus_tag="RP11-418P12.1-003"
 /note="match: cDNAs: ALJ37605.1"
 complement(99959)
 /locus_tag="RP11-418P12.1-003"
 complement(99977..99982)
 /locus_tag="RP11-418P12.1-003"
 join(complement(AL590290.5:2104..2226),
 complement(129944..130027),complement(129613..129720),
 complement(101922..102080))
 /locus_tag="RP11-418P12.1-003"
 /standard_name="OTTHUMP0000016649"
 /note="match: proteins: Q9NT06"
 /codon_start=1
 /protein_id="CAI40661.1"
 /db_xref="GI:57209294"
 /db_xref="GOA:Q9NT06"
 /db_xref="InterPro:IPR000891"
 /translation="AAPGSANGVPSAVKHCLSYQOLLREHLWIGDSVAGALDPAQET
 SQLGLPEFKIVEGPRDGLQNEKVIPTDIEKIEFINRLSQTGLSIEVTSFVSRRW
 VPQVAAAGATEISVFGAASESFSKKNINCISIESMGKPEEVVKSARHNNIPARGYL"
 127828..127923
 /note="Sequence from uni-directional primer reads and dGTP
 big dye terminator reads only."
 145034
 /note="Clone_left_end: RP11-22806"
 164831
 /note="Clone_right_end: RP11-418P12"
 ORIGIN
 Query Match 76.8%; Score 19.2; DB 8; Length 164831;
 Best Local Similarity 87.5%; Pred. No. 1.8e+02;
 Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 2 AAAAAACCTTGGAGGGAGACTCA 25
 |||||
 Db 154107 AAAAAACATTTGGAGGAAAACCTCA 154130
 Search completed: February 3, 2006, 21:26:27
 Job time : 876 secs

GenCore version 5.1.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 3, 2006, 10:53:33 ; Search time 203.111 Seconds
(without alignments)
820.326 Million cell updates/sec

Title: US-10-719-900-10

Perfect score: 25

Sequence: 1 aaaaacacttgaaggagactca 25

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database :

N Geneseq_21.*
1: Geneseqn1980s.*
2: Geneseqn1990s.*
3: Geneseqn2000s.*
4: Geneseqn2001as.*
5: Geneseqn2001bs.*
6: Geneseqn2002as.*
7: Geneseqn2002bs.*
8: Geneseqn2003as.*
9: Geneseqn2003bs.*
10: Geneseqn2003cs.*
11: Geneseqn2003ds.*
12: Geneseqn2004as.*
13: Geneseqn2004bs.*
14: Geneseqn2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	19.2	76.8	873	10	ABT42203
2	19.2	76.8	972	8	ABZ51779
3	18.6	74.4	2511	6	ABK34598 Human CDN
4	18.6	74.4	9210	4	AAL37257 Human mus
5	18.6	74.4	9210	4	AAL37270 Human mus
6	18.6	74.4	9210	8	ABX60245 cDNA enco
7	18.6	74.4	9210	8	ABX60258 cDNA enco
8	18.6	74.4	9210	12	ADJ30995 Human mus
9	18.6	74.4	9210	12	ADJ31008 Human mus
10	18.6	74.4	10818	12	ADQ97332 Human can
11	18.6	74.4	23643	12	ADQ97331 5
12	18.6	74.4	62169	12	ADQ59521 Human can
13	18.6	74.4	63529	14	ADZ13902 Murine ca
14	18.6	74.4	110000	14	ADZ42274_2
15	18.2	72.8	188	12	ACH99910 Human gen
16	18.2	72.8	786	6	ACH76210 Human gen
17	18.2	72.8	786	6	ABZ15673 Arabidops
18	18.2	72.8	865	11	ACN86065 Breast ca
19	18.2	72.8	1260	13	ADS57832 Bacterial

20	18.2	72.8	2092	6	ABL68555
21	18.2	72.8	2092	11	ADI31610
22	18.2	72.8	2092	13	ADS83677 Human lym
23	18.2	72.8	2306	13	ADT47914 Bacterial
24	18.2	72.8	2307	8	ACA57281 Human adi
25	18.2	72.8	6721	12	ADO31133 N-end rul
26	18.2	72.8	6721	13	ADX97505 Pancreat
27	18.2	72.8	6787	14	ADZ49257 Insulin s
28	18.2	72.8	145985	12	ADQ97164 Human can
29	17.8	71.2	346	10	ADD17061 DNA (Segi
30	17.8	71.2	524	10	ADC76449 DNA homol
31	17.8	71.2	524	10	ADC76449 DNA homol
32	17.8	71.2	525	10	ADK56966 Plant DNA
33	17.8	71.2	530	10	ADC75850 DNA homol
34	17.8	71.2	5347	10	ADB75375 Prostata
35	17.8	71.2	30642	13	ABD33003 Mouse can
36	17.8	71.2	72352	12	ADQ97067 Mouse can
37	17.8	71.2	169998	6	ADJ36511 Human Her
38	17.8	71.2	197496	6	ABN85584 Human EGF
39	17.8	71.2	209083	13	ABD332854 Human can
40	17.8	71.2	334462	10	ADC24763 Human wil
41	17.6	70.4	120	3	AAC13345 Human sec
42	17.6	70.4	299	2	AAV04596 Flea seri
43	17.6	70.4	299	4	AAC90858 Flea seri
44	17.6	70.4	300	3	AAA000406 Human col
45	17.6	70.4	389	5	AAF65920 Novel hum
46	17.6	70.4	398	4	AAI88689 Human pol
47	17.6	70.4	436	2	AAV04595 Flea seri
48	17.6	70.4	436	2	AAC90876 Flea seri
49	17.6	70.4	436	5	AAF83431 Flea seri
50	17.6	70.4	480	14	ACL63156 Human col
51	17.6	70.4	555	11	ADT97028 Colon can
52	17.6	70.4	555	11	ADX43510 Human CDN
53	17.6	70.4	563	6	ABQ58300 Human col
54	17.6	70.4	567	6	ABV88555 Human col
55	17.6	70.4	567	3	AAC06210 Human sec
56	17.6	70.4	786	10	ABT41810 Toxicity
57	17.6	70.4	1110	6	ABL91321 Chlamydia
58	17.6	70.4	1152	5	AAF83435 Flea seri
59	17.6	70.4	1152	5	AAF83436 Flea seri
60	17.6	70.4	1303	5	AAF83434 Flea seri
61	17.6	70.4	1303	5	AAF83433 Flea seri
62	17.6	70.4	1607	5	ABV23379 Human pro
63	17.6	70.4	1966	9	ACH04153 Human CDN
64	17.6	70.4	2273	12	ADQ86610 Human tum
65	17.6	70.4	2477	10	ADA53176 Human cod
66	17.6	70.4	2516	10	ADA53332 Human cod
67	17.6	70.4	2588	10	ADC30725 Human nov
68	17.6	70.4	2627	11	ADI30862 Human CDN
69	17.6	70.4	2627	13	ADS82929 Human lym
70	17.6	70.4	3760	6	ABK83731 Human CDN
71	17.6	70.4	3856	10	ADJ56310 Human col
72	17.6	70.4	4015	4	AAH33952 Human pol
73	17.6	70.4	4029	4	AAI58323 Human pol
74	17.6	70.4	4029	5	ADQ98531 DNA enco
75	17.6	70.4	4029	9	ADB48291 Novel hum
76	17.6	70.4	4331	4	AAI60109 Human pol
77	17.6	70.4	5184	4	AAI03097 Human rep
78	17.6	70.4	4061	12	ADQ97540 Mouse can
79	17.6	70.4	94752	10	ADF11646 Human chr
80	17.6	70.4	110000	2	AAK91990_10
81	17.6	70.4	110000	6	ABA90521_00
82	17.6	70.4	140036	6	AAK98600 Human gen
83	17.6	70.4	17472	14	ADZ13139 Human can
84	17.6	70.4	174703	11	ACN44738 Human can
85	17.6	70.4	247544	12	ADQ59419 Human can
86	17.6	70.4	247654	12	ADZ13712 Murine ca
87	17.6	70.4	312477	14	ADP69744 Human ROC
88	17.4	69.6	2682	6	ABZ14037 Arabidops
89	17.4	69.6	109559	13	ABD33157 Murine ca
90	17.2	68.8	300	2	AAZ14942 Human gen
91	17.2	68.8	402	5	AAF66064 Novel hum
92	17.2	68.8	493	5	ABV07171 Human pro

93	17.2	68.8	1647	12	ADN72128
94	17.2	68.8	1647	12	ADN74846
C 95	17.2	68.8	1727	4	AAS26637
C 96	17.2	68.8	1727	8	ABX73986
97	17.2	68.8	1737	10	ADB58591
98	17.2	68.8	1737	10	ADB53224
99	17.2	68.8	1737	13	ADV41392
C 100	17.2	68.8	2775	13	AQO80686
C 101	17.2	68.8	2793	6	ABZ12589
C 102	17.2	68.8	2793	10	ADZ31434
C 103	17.2	68.8	2793	12	ADI44098
104	17.2	68.8	3097	12	ADO35543
105	17.2	68.8	10368	4	ABU05200
C 106	17.2	68.8	42863	10	ADC86412
C 107	17.2	68.8	95484	12	ADQ97298
C 108	17.2	68.8	110000	11	ACN44932_0
C 109	17.2	68.8	110000	14	AEA61124_1
C 110	17	68.0	104	9	ACH31000
C 111	17	68.0	112	9	ACH31003
112	17	68.0	201	13	ADQ41416
113	17	68.0	201	13	ADQ41424
114	17	68.0	201	13	ADQ41414
115	17	68.0	201	13	ADQ41426
116	17	68.0	387	12	ADP91769
117	17	68.0	396	12	ADP91770
118	17	68.0	441	13	ACN52012
119	17	68.0	467	4	AAK63939
120	17	68.0	492	10	ADF14124
C 121	17	68.0	497	9	ACH18320
C 122	17	68.0	524	13	ADQ54154
C 123	17	68.0	547	3	AC78359
C 124	17	68.0	553	3	ACH70372
125	17	68.0	586	3	AAK74925
C 126	17	68.0	591	10	ADD48781
127	17	68.0	603	13	ACN51185
C 128	17	68.0	604	13	ACN53437
C 129	17	68.0	715	10	ADC72291
C 130	17	68.0	982	13	ADX54334
C 131	17	68.0	1095	2	AAV23696
C 132	17	68.0	1104	5	AAS78211
C 133	17	68.0	1104	5	AAS75633
134	17	68.0	1104	5	AAS84590
135	17	68.0	1108	10	ADF14125
C 136	17	68.0	1122	12	ADM36165
137	17	68.0	1137	11	ACL28459
138	17	68.0	1139	4	AAK86252
139	17	68.0	1230	4	ABA89350
C 140	17	68.0	1325	13	ADR24261
C 141	17	68.0	1330	4	AAH34853
142	17	68.0	1530	14	AEA39250
C 143	17	68.0	1575	9	ADB09307
C 144	17	68.0	1575	9	ADB09301
C 145	17	68.0	1575	9	ADB09309
C 146	17	68.0	1575	9	ADB09303
C 147	17	68.0	1575	9	ADB09305
C 148	17	68.0	1636	6	ABK34741
C 149	17	68.0	1656	2	AAK60810
C 150	17	68.0	1656	4	AAS59216

ALIGNMENTS

RESULT 1
ABT42203
ID ABT42203 standard; DNA; 873 BP.
XX
AC ABT42203;
XX
DT 26-JUN-2003 (first entry)
XX
DE Toxicity modelling related rat gene SEQ ID No 1905.
XX

KW Adn72128 Thale cre
KW Adn74846 Thale cre
XX Aas26637 Human gen
OS Abx73986 Human nov
XX Adb58591 Toxicity-
XX Adb53224 Primary r
XX Adv41392 Rat cardi
PD Adq80686 Arabidops
XX Abz12589 Arabidops
XX Ade31434 Plant yie
XX Adi44098 Plant tra
XX Adb35543 Novel mou
PR Abi05200 Drosophi
PR Adc86412 Human GPC
PR Adq97298 Mouse can
PR Acn44932 Mouse gen
PR Continuation (2 of
PR Achn1000 Human bon
PR Achn1003 Human bon
PR Adq41416 Myocardia
PR Adq41424 Myocardia
PR Adq41414 Myocardia
PR Adq41426 Myocardia
PR Adp91769 Cotton ex
PR Adp91770 Cotton ex
PR Acn52012 Cotton an
PR Aak63939 Human inm
PR Adf14124 Human end
PR Achn1320 Human adu
PR Adq54154 Novel can
PR Aac78359 Human can
PR Ach70372 Human gen
PR Aac74925 Human ORF
XX Add48781 Rat gene
XX Acn51185 Cotton an
XX Acn53437 Cotton an
XX Adc72291 Poppy DNA
XX Adx54334 Plant ful
XX Aav23696 Human NSP
XX Aas78211 DNA encod
XX Aas75633 DNA encod
XX Aas84590 DNA encod
XX Adf14125 Human end
XX Adm36165 Human RTN
XX Acl28459 Rice abio
XX Aak86252 Human inm
XX Aab89350 Escherich
XX Aat24261 Breast ca
XX Aah34853 Human col
XX Aea39250 Zizania (
XX Adb09307 Alloioococ
XX Adb09301 Alloioococ
XX Adb09309 Alloioococ
XX Adb09303 Alloioococ
XX Adb09305 Alloioococ
XX Abk34741 Human CDN
XX Aas60810 Human sec
XX Aas59216 Human CDN
SQ Sequence 873 BP; 213 A; 245 C; 225 G; 190 T; 0 U; 0 Other;
Query Match 76.8%; Score 19.2; DB 10; Length 873;
Best Local Similarity 87.5%; Pred. No. 1e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 2 AAAAACTTGGAGGGAGACTCA 25
DB 658 AAGACACTTGCAGGGTGACTCA 681
Toxic effect; gene expression profile; renal toxicity; toxicity marker;
database; drug screening; toxicity assay; rat; ds.
Rattus norvegicus.
WO200295000-A2.
28-NOV-2002.
22-MAY-2002; 2003WO-US016173.
22-MAY-2001; 2001US-0292335P.
13-JUN-2001; 2001US-0297523P.
19-JUN-2001; 2001US-0298925P.
10-JUL-2001; 2001US-0303807P.
10-JUL-2001; 2001US-0303808P.
10-JUL-2001; 2001US-0303810P.
28-AUG-2001; 2001US-0315047P.
27-SEP-2001; 2001US-0324928P.
22-OCT-2001; 2001US-0330462P.
01-NOV-2001; 2001US-0330867P.
21-NOV-2001; 2001US-0331805P.
06-DEC-2001; 2001US-0336144P.
19-DEC-2001; 2001US-0340873P.
21-FEB-2002; 2002US-0357842P.
21-FEB-2002; 2002US-0357843P.
21-FEB-2002; 2002US-0357844P.
15-MAR-2002; 2002US-0364134P.
08-APR-2002; 2002US-0370144P.
08-APR-2002; 2002US-0370206P.
08-APR-2002; 2002US-0370247P.
17-APR-2002; 2002US-0372794P.
21-APR-2002; 2002US-0371679P.
(GENE-) GENE LOGIC INC.
Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Elashoff M;
WPI; 2003-148464/14.
Predicting at least one toxic effect of a compound, useful for toxicity modeling, comprises preparing a gene expression profile of a tissue or cell sample exposed to the compound, and comparing the gene expression profile to a database.
Example 4; Page; 446pp; English.
The invention relates to a novel method of predicting at least one toxic effect of a compound. The method comprises a gene expression profile of a tissue or cell sample exposed to the compound, and comparing the gene expression profile to a database comprising at least part of the data or information given in the specification. The methods are useful for predicting at least one toxic effect of a compound, predicting the renal progression of a toxic effect of a compound, predicting the renal toxicity of a compound, or identifying toxicity markers in tissues or cells exposed to known renal toxin. The genes are useful as toxicity markers in drug screening and toxicity assays, in monitoring disease or physiological states, or disease progression. This polynucleotide represents a rat DNA sequence relating to the toxic effect database described in the specification. NOTE: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from the World Intellectual Property Organization

XX Human musculoskeletal system related polynucleotide SEQ ID NO 3622.
DE
XX
KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;
KW vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; human; secreted protein;
KW musculoskeletal system; ds.
XX
OS Homo sapiens.
XX
PN WO200155367-A1.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001338.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237040P.
PR 02-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-02556719P.


```
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-451937/48.
XX
XX Isolated polypeptide for treating, preventing and/ or prognosing
PT disorders related to the musculoskeletal system including musculoskeletal
PT cancers and also for testing and detection e.g. diagnosis.
XX
XX Example 2; SEQ ID NO 3622; 781pp + Sequence Listing; English.
XX
XX The invention relates to novel genes (AAL34669-AAL37666) and proteins
CC (ABB03087-ABB04109) associated with the musculoskeletal system useful for
CC preventing, treating or ameliorating medical conditions e.g. by protein
CC or gene therapy. The genes are isolated from a range of human tissues
CC disclosed in the specification. The nucleic acids, proteins, antibodies
CC and (ant)agonists are useful in the diagnosis, treatment and prevention
CC of: (a) cancer, e.g. breast and ovarian cancer and other cancers of the
CC adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver,
CC lung, or urogenital; (b) immune disorders e.g. Addison's disease,
CC allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
CC arthritis and ulcerative colitis; (c) cardiovascular disorders such as
CC myocardial ischaemia; (d) wound healing; (e) neurological diseases e.g.
CC cerebral anoxia and epilepsy; and (f) infectious diseases such as viral,
CC bacterial, fungal and parasitic infections. Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 9210 BP; 2878 A; 1772 C; 1716 G; 2844 T; 0 U; 0 Other;
Query Match 74.4%; Score 18.6; DB 4; Length 9210;
Best Local Similarity 84.0%; Pred. No. 2.7e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 AAAAAACACTTGGGAAGGAGACTCA 25
Db 7719 AAACACATTTGGGAAGGAGATTCA 7695
||||| ||||| ||||| ||||| |||||
RESULT 5
AAL37270/c
ID AAL37270 Standard; DNA; 9210 BP.
XX
AC AAL37270;
XX
DT 08-JAN-2002 (first entry)
XX
XX Human musculoskeletal system related polynucleotide SEQ ID NO 3635.
XX
XX Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW anti-allergic; hepatotropic; antidiabetic; anti-inflammatory; antitumor;
KW vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; human; secreted protein;
KW musculoskeletal system; ds.
XX
XX Homo sapiens.
XX
XX WO200155367-A1.
XX
XX 02-AUG-2001.
XX
FD
```

```
XX
PF 17-JAN-2001; 2001WO-US001338.
XX
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 21-SEP-2000; 2000US-0233965P.
PR 21-SEP-2000; 2000US-0234223P.
PR 25-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
```

PR 29-SEP-2000; 2000US-0236369P.
 PR 29-SEP-2000; 2000US-0236370P.
 PR 02-OCT-2000; 2000US-0236802P.
 PR 02-OCT-2000; 2000US-0237037P.
 PR 02-OCT-2000; 2000US-0237038P.
 PR 02-OCT-2000; 2000US-0237039P.
 PR 02-OCT-2000; 2000US-0237040P.
 PR 13-OCT-2000; 2000US-0239935P.
 PR 13-OCT-2000; 2000US-0239937P.
 PR 20-OCT-2000; 2000US-0240960P.
 PR 20-OCT-2000; 2000US-0241221P.
 PR 20-OCT-2000; 2000US-0241785P.
 PR 20-OCT-2000; 2000US-0241786P.
 PR 20-OCT-2000; 2000US-0241787P.
 PR 20-OCT-2000; 2000US-0241808P.
 PR 20-OCT-2000; 2000US-0241809P.
 PR 20-OCT-2000; 2000US-0241826P.
 PR 01-NOV-2000; 2000US-0244617P.
 PR 08-NOV-2000; 2000US-0246474P.
 PR 08-NOV-2000; 2000US-0246475P.
 PR 08-NOV-2000; 2000US-0246476P.
 PR 08-NOV-2000; 2000US-0246477P.
 PR 08-NOV-2000; 2000US-0246478P.
 PR 08-NOV-2000; 2000US-0246523P.
 PR 08-NOV-2000; 2000US-0246524P.
 PR 08-NOV-2000; 2000US-0246525P.
 PR 08-NOV-2000; 2000US-0246526P.
 PR 08-NOV-2000; 2000US-0246527P.
 PR 08-NOV-2000; 2000US-0246528P.
 PR 08-NOV-2000; 2000US-0246532P.
 PR 08-NOV-2000; 2000US-0246609P.
 PR 08-NOV-2000; 2000US-0246610P.
 PR 08-NOV-2000; 2000US-0246611P.
 PR 08-NOV-2000; 2000US-0246613P.
 PR 17-NOV-2000; 2000US-0249207P.
 PR 17-NOV-2000; 2000US-0249208P.
 PR 17-NOV-2000; 2000US-0249209P.
 PR 17-NOV-2000; 2000US-0249210P.
 PR 17-NOV-2000; 2000US-0249211P.
 PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249244P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249264P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 05-DEC-2000; 2000US-0256719P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 08-DEC-2000; 2000US-0251990P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA
 XX Rosen CA, Barash SC, Ruben SM;
 PI
 XX WPI; 2001-451937/48.
 DR
 XX

PT Isolated polypeptide for treating, preventing and/ or prognosing
 PT disorders related to the musculoskeletal system including musculoskeletal
 PT cancers and also for testing and detection e.g. diagnosis.
 XX
 XX Example 2; SEQ ID NO 3635; 781pp + Sequence Listing; English.
 XX
 CC The invention relates to novel genes (AAL34669-AAL37666) and proteins
 CC (ABR03087-ABR04109) associated with the musculoskeletal system useful for
 CC preventing, treating or ameliorating medical conditions e.g. by protein
 CC or gene therapy. The genes are isolated from a range of human tissues
 CC disclosed in the specification. The nucleic acids, proteins, antibodies
 CC and (ant)agonists are useful in the diagnosis, treatment and prevention
 CC of: (a) cancer, e.g. breast and ovarian cancer and other cancers of the
 CC adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver,
 CC lung, or urogenital; (b) immune disorders e.g. Addison's disease,
 CC allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
 CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
 CC arthritis and ulcerative colitis; (c) cardiovascular disorders such as
 CC myocardial ischaemias; (d) wound healing; (e) neurological diseases e.g.
 CC cerebral anoxia and epilepsy; and (f) infectious diseases such as viral,
 CC bacterial, fungal and parasitic infections. Note: The sequence data for
 CC this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 9210 BP; 2878 A; 1772 C; 1716 G; 2844 T; 0 U; 0 Other;
 Query Match 74.4%; Score 18.6; DB 4; Length 9210;
 Best Local Similarity 84.0%; Pred. No. 2.7e-02;
 Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 1 AAAAACAACCTTGGAGGAGGACTCA 25
 Db 7719 AAACACATTTGGAGGAAGATTCA 7695
 RESULT 6
 ABX60245/c
 ID ABX60245 standard; cDNA; 9210 BP.
 XX
 AC ABX60245;
 XX
 DT 26-FEB-2003 (first entry)
 XX
 DE cDNA encoding novel human musculoskeletal system antigen #2589.
 XX
 KW Gene; ss; musculoskeletal system antigen; cancer; metastasis;
 KW re-vascularisation; thrombosis; arteriosclerosis; mineral content;
 KW cardiovascular condition; wound; injury; burn; angiogenesis; ulcer;
 KW post-operative tissue repair; limb regeneration; neuronal growth;
 KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
 KW AIDS-related complex; chondrocyte growth; bone regeneration;
 KW periodontal regeneration; tissue transport; bone graft; skin aging;
 KW keratinocyte growth; hair loss; melanocyte growth; cell proliferation;
 KW cell growth; organ transplant; cell differentiation; body height; weight;
 KW hair colour; eye colour; skin; percentage of adipose tissue;
 KW pigmentation; cosmetic surgery; metabolism; biorhythm; cardiac rhythm;
 KW depression; tendency for violence; pain; reproductive capability;
 KW hormone level; endocrine level; appetite; libido; memory; stress;
 KW storage capability; fat content; lipid content; protein content;
 KW carbohydrate content; vitamin content; cofactor content;
 KW nutritional component.
 XX
 XX Homo sapiens.
 OS
 XX
 XX US2002147140-A1.
 PN
 XX
 XX 10-OCT-2002.
 PD
 XX
 XX 17-JAN-2001; 2001US-00764877.
 PF
 XX
 XX 31-JAN-2000; 2000US-0179065P.
 PR
 PR 04-FEB-2000; 2000US-0180628P.

07-JUL-2000; 2000US-0216647P.
07-JUL-2000; 2000US-0216880P.
11-JUL-2000; 2000US-0217487P.
11-JUL-2000; 2000US-0217496P.
14-JUL-2000; 2000US-0218290P.
26-JUL-2000; 2000US-0220963P.
26-JUL-2000; 2000US-0220964P.
14-AUG-2000; 2000US-0224518P.
14-AUG-2000; 2000US-0224519P.
14-AUG-2000; 2000US-0225267P.
14-AUG-2000; 2000US-0225268P.
14-AUG-2000; 2000US-0225270P.
14-AUG-2000; 2000US-0225447P.
14-AUG-2000; 2000US-0225757P.
14-AUG-2000; 2000US-0225758P.
22-AUG-2000; 2000US-0226868P.
30-AUG-2000; 2000US-0228924P.
01-SEP-2000; 2000US-0229287P.
01-SEP-2000; 2000US-0229343P.
01-SEP-2000; 2000US-0229344P.
01-SEP-2000; 2000US-0229345P.
05-SEP-2000; 2000US-0229509P.
08-SEP-2000; 2000US-0229513P.
08-SEP-2000; 2000US-0231413P.
21-SEP-2000; 2000US-0234223P.
21-SEP-2000; 2000US-0234274P.
25-SEP-2000; 2000US-0234997P.
27-SEP-2000; 2000US-0235834P.
29-SEP-2000; 2000US-0236327P.
29-SEP-2000; 2000US-0236367P.
29-SEP-2000; 2000US-0236368P.
29-SEP-2000; 2000US-0236369P.
29-SEP-2000; 2000US-0236370P.
02-OCT-2000; 2000US-0236802P.
02-OCT-2000; 2000US-0237037P.
02-OCT-2000; 2000US-0237038P.
02-OCT-2000; 2000US-0237039P.
02-OCT-2000; 2000US-0237040P.
13-OCT-2000; 2000US-0239353P.
20-OCT-2000; 2000US-0240960P.
20-OCT-2000; 2000US-0241785P.
20-OCT-2000; 2000US-0241809P.
01-NOV-2000; 2000US-0244617P.
17-NOV-2000; 2000US-0249299P.
08-DEC-2000; 2000US-0251856P.
08-DEC-2000; 2000US-0251868P.
08-DEC-2000; 2000US-0251869P.

(ROSE/) ROSEN C A.
(RUBE/) RUBEN S M.
(BARA/) BARASH S C.

Rosen CA, Ruben SM, Barash SC;
WPI; 2003-128199/12.

Isolated nucleic acid molecules encoding musculoskeletal system associated polypeptides, useful for detecting disorders, e.g. cancer.

Disclosure; SEQ ID NO 3635; 321pp; English.

The invention describes an isolated nucleic acid molecule comprising a sequence encoding musculoskeletal system associated polypeptides useful for detecting disorders, e.g., cancer or cancer metastases, in animals or humans. The nucleic acid; stimulates re-vascularisation of ischaemic tissues associated with conditions such as thrombosis, arteriosclerosis, and other cardiovascular conditions; treats wounds due to injuries, burns, post-operative tissue repair, and ulcers; stimulates angiogenesis and limb regeneration; stimulates neuronal growth; can treat and prevent neuronal damage occurring in certain disorders or neurodegenerative conditions, such as, Alzheimer's disease, Parkinson's disease, and AIDS-related complex; stimulates chondrocyte growth, thus they can be used to enhance bone and periodontal regeneration and aid in tissue transports or

CC bone grafts; prevents skin aging due to sunburn by stimulating
CC keratinocyte growth; prevents hair loss, since FGF family members
CC activate hair-forming cells and promotes melanocyte growth; stimulates
CC growth and differentiation of hematopoietic cells and bone marrow cells
CC when used in combination with other cytokines; maintains organs before
CC transplantation or for supporting cell culture of primary tissues;
CC induces tissue of mesodermal origin to differentiate in early embryos;
CC increases or decreases the differentiation or proliferation of embryonic
CC stem cells, besides, haematopoietic lineage; modulates mammalian
CC characteristics, such as, body height, weight, hair colour, eye colour,
CC skin, percentage of adipose tissue, pigmentation, size, and shape (e.g.,
CC cosmetic surgery); modulates mammalian metabolism; changes mammal's metal
CC state or physical state by influencing biorhythms, circadian rhythms,
CC depression, tendency for violence, tolerance for pain, reproductive
CC capabilities, hormonal or endocrine levels, appetite, libido, memory, or
CC stress; increases or decreases storage capabilities, fat content, lipid,
CC protein, carbohydrate, vitamins, minerals, cofactors or other nutritional
CC components. This sequence encodes a novel human musculoskeletal system
CC antigen. Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from the US patent office at
CC ftp.segdata.uspto.gov/sequence.html?DocID=20020147140
XX
SQ Sequence 9210 BP; 2878 A; 1772 C; 1716 G; 2844 T; 0 U; 0 Other;

Query Match 74.4%; Score 18.6; DB 8; Length 9210;
Best Local Similarity 84.0%; Pred. No. 2.7e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAACACTTGGAGGAGACTCA 25
||||| ||||| ||||| |||||
Db 7719 AAACACATTGGAGGAAGATTCA 7695

RESULT 8
ADJ30995/c
ID ADJ30995 standard; DNA; 9210 BP.
XX
AC ADJ30995;
XX
DT 20-MAY-2004 (first entry)
XX
DE Human musculoskeletal system-associated genomic DNA - SEQ ID 3622.
XX
KW musculoskeletal system; cytostatic; osteopathic; cancer; osteoporosis;
KW gene therapy; vaccine; human; ds.
XX
OS Homo sapiens.
XX
PN US2004009488-A1.
XX
PD 15-JAN-2004.
XX
PF 13-SEP-2002; 2002US-00242515.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.

PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-022547P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226661P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 22-AUG-2000; 2000US-0227009P.
PR 23-AUG-2000; 2000US-0227099P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 06-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249267P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 03-DEC-2000; 2000US-0251030P.
PR 03-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
PR 17-JAN-2001; 2001US-00764877.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Ruben SM, Barash SC;
XX WFI; 2004-090458/09.
XX
XX New nucleic acid molecule, useful for preparing a medicament for
PT preventing, treating or ameliorating a medical condition e.g., cancer of
PT musculoskeletal tissues or osteoporosis.
XX
XX Disclosure; SEQ ID NO 3622; 289pp; English.
PS
XX The invention relates to a novel isolated musculoskeletal system-
XX associated nucleic acid molecule. The nucleic acid of the invention
CC demonstrates cytostatic and osteopathic activities and may be useful for
CC preparing a medicament for preventing, treating or ameliorating a medical
CC condition such as cancer of the musculoskeletal tissues or osteoporosis,
CC possibly via gene therapy or vaccine production. The current sequence is
CC that of the human musculoskeletal system-associated genomic DNA of the
CC invention. The current sequence is not shown within the specification per
CC se but is available on the USPTO web-site
CC http://seqdata.uspto.gov/sequence.html?DocID=20040009488.
XX
XX Sequence 9210 BP; 2878 A; 1772 C; 1716 G; 2844 T; 0 U; 0 Other;
SQ Query Match 74.4%; Score 18.6; DB 12; Length 9210;

Best Local Similarity 84.0%; Pred. No. 2.7e+02;			
Matches	21; Conservative	0; Mismatches	4; Indels
		0; Gaps	0;
QY	1 AAAAAACACTTGGAGGAGACTCA 25		
Db	7719 AAAACACATTGGAAGGAGATTCA 7695		
RESULT 9			
ID	ADJ31008/c		
ADJ31008	standard; DNA; 9210 BP.		
XX	AC ADJ31008;		
XX			
DT	20-MAY-2004 (first entry)		
XX			
DE	Human musculoskeletal system-associated genomic DNA - SEQ ID 3635.		
XX			
KW	musculoskeletal system; cytostatic; osteopathic; cancer; osteoporosis;		
KW	gene therapy; vaccine; human; ds.		
XX			
OS	Homo sapiens.		
XX			
PN	US2004009488-A1.		
XX			
PD	15-JAN-2004.		
XX			
PF	13-SEP-2002; 2002US-00242515.		
XX			
PR	31-JAN-2000; 2000US-0179065P.		
PR	04-FEB-2000; 2000US-0180628P.		
PR	24-FEB-2000; 2000US-0184664P.		
PR	02-MAR-2000; 2000US-0186350P.		
PR	16-MAR-2000; 2000US-0189874P.		
PR	17-MAR-2000; 2000US-0190076P.		
PR	18-APR-2000; 2000US-0198123P.		
PR	19-MAY-2000; 2000US-0205515P.		
PR	07-JUN-2000; 2000US-0209467P.		
PR	28-JUN-2000; 2000US-0214886P.		
PR	30-JUN-2000; 2000US-0215135P.		
PR	07-JUL-2000; 2000US-0216647P.		
PR	07-JUL-2000; 2000US-0216880P.		
PR	11-JUL-2000; 2000US-0217487P.		
PR	11-JUL-2000; 2000US-0217496P.		
PR	14-JUL-2000; 2000US-0218290P.		
PR	26-JUL-2000; 2000US-0220963P.		
PR	26-JUL-2000; 2000US-0220964P.		
PR	14-AUG-2000; 2000US-0224518P.		
PR	14-AUG-2000; 2000US-0224519P.		
PR	14-AUG-2000; 2000US-0225213P.		
PR	14-AUG-2000; 2000US-0225214P.		
PR	14-AUG-2000; 2000US-0225266P.		
PR	14-AUG-2000; 2000US-0225267P.		
PR	14-AUG-2000; 2000US-0225268P.		
PR	14-AUG-2000; 2000US-0225270P.		
PR	14-AUG-2000; 2000US-0225447P.		
PR	14-AUG-2000; 2000US-0225757P.		
PR	14-AUG-2000; 2000US-0225758P.		
PR	14-AUG-2000; 2000US-0225759P.		
PR	18-AUG-2000; 2000US-0226279P.		
PR	22-AUG-2000; 2000US-0226681P.		
PR	22-AUG-2000; 2000US-0226868P.		
PR	22-AUG-2000; 2000US-0227182P.		
PR	23-AUG-2000; 2000US-0227009P.		
PR	30-AUG-2000; 2000US-0228924P.		
PR	01-SEP-2000; 2000US-0229287P.		
PR	01-SEP-2000; 2000US-0229343P.		
PR	01-SEP-2000; 2000US-0229344P.		
PR	01-SEP-2000; 2000US-0229345P.		
PR	05-SEP-2000; 2000US-0229509P.		
PR	05-SEP-2000; 2000US-0229513P.		
PR	06-SEP-2000; 2000US-0230437P.		
PR	06-SEP-2000; 2000US-0230438P.		

PR	08-SEP-2000; 2000US-0231242P.
PR	08-SEP-2000; 2000US-0231243P.
PR	08-SEP-2000; 2000US-0231244P.
PR	08-SEP-2000; 2000US-0231413P.
PR	08-SEP-2000; 2000US-0231414P.
PR	08-SEP-2000; 2000US-0232080P.
PR	08-SEP-2000; 2000US-0232081P.
PR	12-SEP-2000; 2000US-0231968P.
PR	14-SEP-2000; 2000US-0232397P.
PR	14-SEP-2000; 2000US-0232398P.
PR	14-SEP-2000; 2000US-0232399P.
PR	14-SEP-2000; 2000US-0232400P.
PR	14-SEP-2000; 2000US-0232401P.
PR	14-SEP-2000; 2000US-0233063P.
PR	14-SEP-2000; 2000US-0233064P.
PR	14-SEP-2000; 2000US-0233065P.
PR	21-SEP-2000; 2000US-0234223P.
PR	21-SEP-2000; 2000US-0234274P.
PR	25-SEP-2000; 2000US-0234997P.
PR	25-SEP-2000; 2000US-0234998P.
PR	26-SEP-2000; 2000US-0235484P.
PR	27-SEP-2000; 2000US-0235834P.
PR	27-SEP-2000; 2000US-0235836P.
PR	29-SEP-2000; 2000US-0236327P.
PR	29-SEP-2000; 2000US-0236367P.
PR	29-SEP-2000; 2000US-0236368P.
PR	29-SEP-2000; 2000US-0236369P.
PR	29-SEP-2000; 2000US-0236370P.
PR	02-OCT-2000; 2000US-0236802P.
PR	02-OCT-2000; 2000US-0237037P.
PR	02-OCT-2000; 2000US-0237038P.
PR	02-OCT-2000; 2000US-0237039P.
PR	02-OCT-2000; 2000US-0237040P.
PR	13-OCT-2000; 2000US-0239933P.
PR	13-OCT-2000; 2000US-0239937P.
PR	20-OCT-2000; 2000US-0240960P.
PR	20-OCT-2000; 2000US-0241221P.
PR	20-OCT-2000; 2000US-0241785P.
PR	20-OCT-2000; 2000US-0241786P.
PR	20-OCT-2000; 2000US-0241787P.
PR	20-OCT-2000; 2000US-0241808P.
PR	20-OCT-2000; 2000US-0241809P.
PR	20-OCT-2000; 2000US-0241826P.
PR	01-NOV-2000; 2000US-0244617P.
PR	08-NOV-2000; 2000US-0246474P.
PR	08-NOV-2000; 2000US-0246475P.
PR	08-NOV-2000; 2000US-0246476P.
PR	08-NOV-2000; 2000US-0246477P.
PR	08-NOV-2000; 2000US-0246478P.
PR	08-NOV-2000; 2000US-0246523P.
PR	08-NOV-2000; 2000US-0246524P.
PR	08-NOV-2000; 2000US-0246525P.
PR	08-NOV-2000; 2000US-0246526P.
PR	08-NOV-2000; 2000US-0246527P.
PR	08-NOV-2000; 2000US-0246528P.
PR	08-NOV-2000; 2000US-0246532P.
PR	08-NOV-2000; 2000US-0246609P.
PR	08-NOV-2000; 2000US-0246610P.
PR	08-NOV-2000; 2000US-0246611P.
PR	08-NOV-2000; 2000US-0246613P.
PR	17-NOV-2000; 2000US-0249207P.
PR	17-NOV-2000; 2000US-0249208P.
PR	17-NOV-2000; 2000US-0249210P.
PR	17-NOV-2000; 2000US-0249211P.
PR	17-NOV-2000; 2000US-0249212P.
PR	17-NOV-2000; 2000US-0249213P.
PR	17-NOV-2000; 2000US-0249214P.
PR	17-NOV-2000; 2000US-0249215P.
PR	17-NOV-2000; 2000US-0249216P.
PR	17-NOV-2000; 2000US-0249217P.
PR	17-NOV-2000; 2000US-0249218P.
PR	17-NOV-2000; 2000US-0249244P.

PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
PR 17-JAN-2001; 2001US-00764877.
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Ruben SM, Barash SC;
PI WPI; 2004-090458/09.
DR
XX
XX New nucleic acid molecule, useful for preparing a medicament for
PT preventing, treating or ameliorating a medical condition e.g., cancer of
PT musculoskeletal tissues or osteoporosis.
XX
XX Disclosure; SEQ ID NO 3635; 289pp; English.
PS
XX
XX The invention relates to a novel isolated musculoskeletal system-
CC associated nucleic acid molecule. The nucleic acid of the invention
CC demonstrates cytostatic and osteopathic activities and may be useful for
CC preparing a medicament for preventing, treating or ameliorating a medical
CC condition such as cancer of the musculoskeletal tissues or osteoporosis,
CC possibly via gene therapy or vaccine production. The current sequence is
CC that of the human musculoskeletal system-associated genomic DNA of the
CC invention. The current sequence is not shown within the specification per
CC se but is available on the USPTO web-site
CC http:seqdata.uspto.gov/sequence.html?DocID=20040009488.
XX
XX Sequence 9210 BP; 2878 A; 1772 C; 1716 G; 2844 T; 0 U; 0 Other;
SQ
Query Match 74.4%; Score 18.6; DB 12; Length 9210;
Best Local Similarity 84.0%; Pred. No. 2.7e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 AAAAAACACTTGGAGGAGACTCA 25
Db 7719 AAAACACATTGGAGGAAGATTCA 7695
RESULT 10
ADQ97332/c
ID ADQ97332 standard; DNA; 10818 BP.
XX
XX AC
XX ADQ97332;
XX
DT 07-OCT-2004 (first entry)
XX
XX Human cancer associated sequence HR08-032, SEQ ID 309.
XX
XX Cytostatic; Gene Therapy; cancer; leukemia; lymphoma; Human; ds.
XX
XX Homo sapiens.
OS
XX WO2004060304-A2.
PN
XX
XX 22-JUL-2004.
PD
XX

PF 22-DEC-2003; 2003WO-US041389.
XX
XX 27-DEC-2002; 2002US-00330773.
XX
XX (SAGR-) SAGRES DISCOVERY INC.
PA
XX Morris DW, Malandro MS;
PI WPI; 2004-543781/52.
XX
XX New isolated cancer associated nucleic acids comprising at least 10
PT contiguous nucleotides, useful for diagnosing, preventing and/or treating
PT cancers such as leukemia and lymphoma.
XX
XX Claim 1; SEQ ID NO 309; 199pp; English.
PS
XX
XX The present invention relates to cancer associated sequences (ADQ97025-
CC ADQ98004). The sequences are useful for the diagnosis, prevention and/or
CC treatment of cancer, such as leukemia and lymphoma. Note: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 10818 BP; 3402 A; 2198 C; 2281 G; 2937 T; 0 U; 0 Other;
SQ
Query Match 74.4%; Score 18.6; DB 12; Length 10818;
Best Local Similarity 84.0%; Pred. No. 2.8e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 AAAAAACACTTGGAGGAGACTCA 25
Db 7262 AAAACACATTGGAGGAAGATTCA 7238
RESULT 11
ADQ97331_5/c
Continuation (6 of 6) of ADQ97331 from base 500001 (Human cancer associated sequence HD
WP Sequence split into 6 fragments LOCUS ADQ97331 Accession Adq97331
WP Fragment Name Begin End
WP ADQ97331_0 1 110000
WP ADQ97331_1 100001 210000
WP ADQ97331_2 200001 310000
WP ADQ97331_3 300001 410000
WP ADQ97331_4 400001 510000
WP ADQ97331_5 500001 523643
Query Match 74.4%; Score 18.6; DB 12; Length 23643;
Best Local Similarity 84.0%; Pred. No. 3.1e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 AAAAAACACTTGGAGGAGACTCA 25
Db 10087 AAAACACATTGGAGGAAGATTCA 10063
RESULT 12
ADQ59521/c
ID ADQ59521 standard; DNA; 62169 BP.
XX
XX AC
XX ADQ59521;
XX
XX 07-OCT-2004 (first entry)
XX
XX Human cancer-associated (CA) gene sequence SEQ ID NO:157.
XX
XX human; cancer-associated gene; cancer-associated protein; cytostatic;
XX gene therapy; vaccine; tyrosine kinase antagonist;
XX G-protein coupled receptor antagonist; cancer; lymphoma; gene; ds.
XX
XX Homo sapiens.
OS
XX WO2004058288-A1.
PN
XX

PD 15-JUL-2004.
XX
XX
XX 15-DEC-2003; 2003WO-US040082.
XX
XX 17-DEC-2002; 2002US-00322696.
XX (SAGR-) SAGRES DISCOVERY INC.
XX
XX Morris DW, Malandro MS;
XX
XX WPI; 2004-543349/52.
DR P-PSDB; ADQ59523.
XX
XX New cancer-associated nucleic acid for diagnosing, preventing or treating
PT cancer (e.g. lymphoma) or for screening agents that may be used for
PT treating or preventing cancer.
XX
XX Disclosure; SEQ ID NO 157; 143pp; English.
XX
XX The present invention describes human cancer-associated (CA) nucleotide
CC sequences (I). Also described: (1) an expression vector comprising (I);
CC (2) a host cell comprising (I) or the expression vector; (3) a microarray
CC for detecting a CA nucleic acid; (4) an isolated polypeptide encoded
CC within an open reading frame of a CA sequence; (5) an isolated antibody,
CC or its antigen binding fragment, that binds to the above polypeptide; (6)
CC a hybridoma that produces the monoclonal antibody described above; (7) a
CC pharmaceutical composition comprising the antibody and a pharmaceutical
CC excipient; (8) a kit for detecting or diagnosing cancer cells, comprising
CC the above (monoclonal) antibody or polynucleotide that selectively
CC hybridizes to any of the polynucleotide sequences mentioned above; (9)
CC methods for diagnosing cancer or for detecting the presence or absence of
CC cancer cells in an individual; (10) a method for inhibiting growth of
CC cancer cells in an individual; (11) a method for delivering a therapeutic
CC agent to cancer cells in an individual; (12) an electronic library
CC comprising the polynucleotide or polypeptide, or their fragments,
CC mentioned above; (13) a method of screening for anticancer activity; (14)
CC methods for detecting cancer associated with expression of a polypeptide
CC or the presence of the antibody in a test cell or serum sample; (15) a
CC method for screening for a bioactive agent capable of modulating the
CC activity of a CA protein encoded by the above nucleic acid molecule; and
CC (16) a method for treating cancers. (I) has cytostatic activity, and can
CC be used in gene therapy, in vaccines, as a tyrosine kinase antagonist,
CC and as a G-protein coupled receptor antagonist. The compositions and
CC methods of the present invention can be used for diagnosing, preventing
CC and treating cancer, especially lymphomas. They may also be used in
CC screening for agents that may be used for treating or preventing cancer.
CC The present sequence represents a human CA gene sequence, which is given
CC in the exemplification of the present invention. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 62169 BP; 17028 A; 13404 C; 13215 G; 18292 T; 0 U; 230 Other;
SQ
Query Match 74.4%; Score 18.6; DB 12; Length 62169;
Best Local Similarity 84.0%; Pred. No. 3.6e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 AAAAAACACTTGAAGGGAGACTCA 25
DB 7740 AAAGGAACCTTGAAGGGTGACTCA 7716
RESULT 13
ID ADZ13902/C
ADZ13902 standard; DNA; 63529 BP.
XX
AC ADZ13902;
XX
XX 16-JUN-2005 (first entry)
XX
XX Murine cancer-associated genomic DNA #124.
DE
XX

KW Diagnosis; DNA microarray; microarray; biochip; cancer; neoplasm;
KW cytostatic; gene; ds.
XX
XX Mus sp.
XX
XX WO2005031001-A2.
XX
XX 07-APR-2005.
XX
XX 23-SEP-2004; 2004WO-US031617.
XX
XX 23-SEP-2003; 2003US-00669920.
XX (CHIR) CHIRON CORP.
XX
XX Morris DW, Malandro MS;
XX
XX WPI; 2005-273395/28.
XX
XX Nucleic acid array useful for detecting cancer associated nucleic acid,
PT comprises two or more nucleic acid probes.
XX
XX Disclosure; SEQ ID NO 1422; 199pp; English.
XX
XX The invention relates to a nucleic acid array for detecting a cancer
CC associated (CA) nucleic acid, comprising two or more nucleic acid probes.
CC The invention also relates to a peptide array comprising two or more
CC isolated polypeptides encoded by a CA nucleic acid sequence, a compound
CC that binds to a polypeptide, an isolated antibody or its fragment which
CC binds to a polypeptide, which is prepared by immunizing a host animal
CC with a composition comprising the polypeptide or its antigen binding
CC fragment and collecting cells from the host expressing antibodies against
CC the antigen or its antigen binding fragment, a composition comprising the
CC antibody and a carrier, a method of screening for anticancer activity, a
CC method of detecting a CA nucleic acid, a method of diagnosing cancer, a
CC method of treating cancer and a method of inhibiting expression of a CA
CC nucleic acid in a cell. The CA nucleic acids are useful for detecting CA
CC nucleic acids. The antibody is useful for detecting the presence or
CC absence of cancer cells in an individual which involves contacting cells
CC from the individual with the antibody and detecting a complex of a CA
CC protein from the cancer cells and the antibody, where the detection of
CC the complex correlates with the presence of cancer cells in the
CC individual. The composition is useful for inhibiting growth of cancer
CC cells in an individual or for delivering a therapeutic agent to cancer
CC cells in an individual. The invention is also useful for diagnosing
CC cancer, for treating cancer and for inhibiting expression of a CA gene in
CC a cell. This sequence represents murine cancer-associated genomic DNA of
CC the invention.
XX
XX Sequence 63529 BP; 17363 A; 13779 C; 13492 G; 18665 T; 0 U; 230 Other;
SQ
Query Match 74.4%; Score 18.6; DB 14; Length 63529;
Best Local Similarity 84.0%; Pred. No. 3.6e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 AAAAAACACTTGAAGGGAGACTCA 25
DB 7763 AAAGGAACCTTGAAGGGTGACTCA 7739
RESULT 14
ADZ42274_2
Continuation (3 of 4) of ADZ42274 from base 200001 (Human mineralocorticoid receptor ge
WP Sequence split into 4 fragments LOCUS ADZ42274 Accession ADZ42274
WP Fragment Name Begin End
WP ADZ42274_0 1 110000
WP ADZ42274_1 100001 210000
WP ADZ42274_2 200001 310000
WP ADZ42274_3 300001 369000
Query Match 74.4%; Score 18.6; DB 14; Length 110000;
Best Local Similarity 84.0%; Pred. No. 3.9e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 AAAAAACACTTGGAGGGGAGACTCA 25
||| ||| ||| ||| ||| ||| ||| |||
Dp 45145 AGAAGACACTTCGAAGGGGAGATCA 45169

RESULT 15	
ACH89910/c	
ID	ACH89910 standard; DNA; 188 BP.
XX	
XX	ACH89910;
XX	
DT	29-JUL-2004 (first entry)
XX	
DE	Human genome derived single exon probe #23105.
XX	
KW	Human; probe; ss; gene expression; single exon probe; microarray;
KW	alternative splicing event; genomic alteration.
XX	
OS	Homo sapiens.
XX	
PN	US2003194704-A1.
XX	
PD	16-OCT-2003.
XX	
PF	03-APR-2002; 2002US-00029386.
XX	
PR	03-APR-2002; 2002US-00029386.
XX	
PA	(PENN/) PENN S G.
PA	(RANK/) RANK D R.
PA	(HANZ/) HANZEL D K.
XX	
PI	Penn SG, Rank DR, Hanzel DK;
XX	
DR	WPI; 2004-119264/12.
XX	
PT	New human genome-derived single exon nucleic acid probes useful for
PT	gene expression analysis, for identifying or characterizing alter
PT	splicing events, for assessing genomic alterations or as tools fo
PT	surveying tissues.

CC	smaller genomic alterations, in priming the synthesis of nucleic acids, or in expressing the ORF-encoded peptide. The present sequence is a human single exon probe of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from USPTO at seqdata.uspto.gov/sequence.html?DocID=20030194704
XX	
SQ	Sequence 188 BP; 54 A; 34 C; 23 G; 77 T; 0 U; 0 Other;
	Query Match 72.8%; Score 18.2; DB 12; Length 188;
	Best Local Similarity 87.0%; Pred. No. 2.4e+02;
	Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Oy	1 AAAAAACACTTGGGAAGGAGACT 23
Db	86 AAACAACACTTGGCAAGGAGACT 64
RESULT 16	
ACH76210/c	
ID	ACH76210 standard; DNA; 590 BP.
XX	
AC	ACH76210;
XX	
DT	29-JUL-2004 (first entry)
XX	
DE	Human genome derived single exon probe #9405.
XX	
KW	Human; probe; ss; gene expression; single exon probe; microarray; alternative splicing event; genomic alteration.
KW	
XX	
OS	Homo sapiens.
XX	
PN	US2003194704-A1.
XX	
PD	16-OCT-2003.
XX	
PF	03-APR-2002; 2002US-00029386.
XX	
PR	03-APR-2002; 2002US-00029386.
XX	
PA	(PENN/) PENN S G.
PA	(RANK/) RANK D R.
PA	(HANSZ/) HANZEL D K.
XX	
PI	Penn SG, Rank DR, Hanzel DK;
XX	
DR	WPI; 2004-119264/12.
XX	
PT	New human genome-derived single exon nucleic acid probes useful for human gene expression analysis, for identifying or characterizing alternative splicing events, for assessing genomic alterations or as tools for surveying tissues.
PT	
XX	
PS	Claim 15; SEQ ID NO 9405; 80pp; English.
XX	
CC	The invention relates to a nucleic acid probe for measuring human gene expression, comprising any of the 27,400 fully defined nucleotide sequences in the specification, or their complements or fragments, and encoding at least 8 amino acids of any of the 6888 amino acid sequences fully defined in the specification. The probe is a single exon probe that hybridises under high stringency conditions to a nucleic acid molecule expressed in human cells or tissues. Also included are a spatially-addressable set of single exon nucleic acid probes for measuring human gene expression (comprising a plurality of single exon nucleic acid probes cited above, where each of the plurality of probes is separately and addressably isolatable or amplifiable from the plurality), a single exon microarray for measuring human gene expression, a method of measuring human gene expression, a vector comprising the single exon probe cited above, an ORF-encoded peptide comprising at least 8 contiguous amino acids of any of the above-mentioned amino acid sequences (optionally with conservative amino acid substitutions), an isolated antibody that binds specifically to a peptide cited above,

CC methods of selling and/or licensing single exon probes or microarrays to
CC a customer desiring to measure gene expression, a method of providing
CC human gene expression data by subscription, and a computer-readable
CC storage medium which contains a database having a plurality of records
CC (each record including data on the expression of a single exon probe
CC cited above. The probe, methods and apparatus are useful in gene
CC expression analysis. The probes may be used as tools for surveying
CC tissues to detect the presence of expressed messages that contain their
CC specific exon, or in constructing genome-derived single exon microarrays.
CC In addition, the probes are used in identifying and characterising
CC alternative splicing events, in detecting and characterising gross
CC alterations in the genomic locus that includes their exon, in assessing
CC smaller genomic alterations, in priming the synthesis of nucleic acids,
CC or in expressing the ORF-encoded peptide. The present sequence is a human
CC single exon probe of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20030194704
XX
SQ Sequence 590 BP; 193 A; 112 C; 81 G; 204 T; 0 U; 0 Other;

Query Match 72.8%; Score 18.2; DB 12; Length 590;
Best Local Similarity 87.0%; Pred. No. 2.8e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAACAACCTTGGAGGGAGACT 23
DB 275 AAACACACTTGGCAAGGAGACT 253
|||||

RESULT 17
ABZ15673
ID ABZ15673 standard; DNA; 786 BP.
XX
AC ABZ15673;
XX
DT 21-JAN-2003 (first entry)
XX
DE Arabidopsis thaliana stress regulated gene SEQ ID NO 3478.
XX
KW Arabidopsis thaliana; plant; gene; stress; transgenic; ds.
XX
OS Arabidopsis thaliana.
XX
PN WO200216655-A2.
XX
PD 28-FEB-2002.
XX
PF 24-AUG-2001; 2001WO-US026685.
XX
PR 24-AUG-2000; 2000US-0227866P.
PR 26-JAN-2001; 2001US-0264647P.
PR 22-JUN-2001; 2001US-0300111P.
XX
PA (SCRI) SCRIPPS RES INST.
PA (SYGN) SYNGENTA PARTICIPATIONS AG.
XX
PI Harper JF, Kreps J, Wang X, Zhu T;
XX
DR WPI; 2002-304127/34.
XX

XX Identifying a stress condition to which a plant cell has been exposed and
PT producing plants with increased tolerance to these abiotic stresses.
XX
PS Claim 144; SEQ ID NO 3478; 577bp + Sequence Listing; English.
XX
CC The invention relates to identifying a stress condition to which a plant
CC cell has been exposed, comprising: (a) contacting nucleic acid
CC representative of expressed polynucleotides in the plant cell with an
CC array or probes representative of the plant cell genome; and (b)
CC detecting a profile of expressed polynucleotides in the plant cell
CC characteristic of a stress response. The method is useful in the
CC production of transgenic plants, cells and seeds and in producing plants

CC with increased tolerance to abiotic stress. The present sequence is that
CC of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used
CC in methods of the invention. Note: The sequence data for this patent is
CC not represented in the printed specification but is based on sequence
CC information supplied to Derwent by the European Patent Office
XX
SQ Sequence 786 BP; 248 A; 147 C; 118 G; 273 T; 0 U; 0 Other;

Query Match 72.8%; Score 18.2; DB 6; Length 786;
Best Local Similarity 87.0%; Pred. No. 2.9e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAACAACCTTGGAGGGAGACTC 24
DB 512 ATAAACAACCTTGGAGAGAGAAATC 534
|||||

RESULT 18
ACN86065/c
ID ACN86065 standard; DNA; 865 BP.
XX
AC ACN86065;
XX
DT 02-DEC-2004 (first entry)
XX
DE Breast cancer related marker, seq id 72115.
XX
KW Cancer; breast; tumour; cytostatic; marker; detection; therapy; ds.
XX
OS Homo sapiens.
XX
PN US2003099974-A1.
XX
PD 29-MAY-2003.
XX
PF 18-JUL-2002; 2002US-00198846.
XX
PR 18-JUL-2001; 2001US-0306220P.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Lillie J, Xu Y, Wang Y, Steinmann K;
XX
DR WPI; 2003-787014/74.
XX
PT Novel isolated polypeptide associated with breast cancer, useful for
PT detecting presence of polypeptide in sample, as a marker for breast
PT cancer.
XX
PS Disclosure; SEQ ID NO 7215; 36pp; English.

XX The invention relates to an isolated polypeptide (I) associated with a
CC breast cancer which is encoded by a nucleic acid molecule comprising a
CC nucleotide sequence (S1). Further disclosed is an antibody that binds to
CC the polypeptide of the invention. The activity of the polypeptide of the
CC invention may be described as cytostatic. The antibody is useful for
CC detecting the presence of (I) in a sample. Nucleic acid molecules of the
CC invention are useful in the detection of breast tumours. (I) is useful as
CC a marker for breast cancer and in breast cancer therapy. Sequences given
CC in records ACN78851-ACN92934 represent nucleic acid markers associated
CC with breast cancer. Note: The sequence listing does not form part of the
CC specification but may be obtained in electronic format from the USPTO web
CC site at seqdata.uspto.gov/sequence.html?DocID=20030099974
XX

SQ Sequence 865 BP; 227 A; 177 C; 174 G; 246 T; 0 U; 41 Other;

Query Match 72.8%; Score 18.2; DB 11; Length 865;
Best Local Similarity 87.0%; Pred. No. 2.9e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAACAACCTTGGAGGGAGACT 23
DB 575 AAAAAGACTTGGAGGTAACACT 553
|||||


```
PR 01-NOV-2000; 2000US-0245084P.
XX
XX (AVAL-) AVALON PHARM.
XX
XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
XX
XX WPI; 2002-188264/24.
XX
XX Screening for anti-neoplastic agent involves exposing cells to a chemical
PT agent to be tested for anti-neoplastic activity, and determining a change
PT in expression of a gene of a signature gene set.
XX
XX Claim 1; SEQ ID NO 6892; 44pp; English.
XX
XX The present invention describes a method (M1) for screening for an anti-
CC neoplastic agent. The method involves exposing cells to a chemical agent
CC to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC to ABL70110), or is at least 9% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening an
CC anti-neoplastic agent, and can be used for producing a product which is
CC the data collected with respect to the anti-neoplastic agent as a result
CC of M1, and the data is sufficient to convey the chemical structure and/or
CC properties of the agent. M1 can be used in the treatment of cancer such
CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,
CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilm's
CC tumour
XX
XX Sequence 2092 BP; 728 A; 359 C; 443 G; 562 T; 0 U; 0 Other;
SQ
Query Match 72.8%; Score 18.2; DB 6; Length 2092;
Best Local Similarity 87.0%; Pred. No. 3.3e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAACAACCTTGGGAAGGAGACT 23
Db 724 AAACAACACTTGGCAAGGAGACT 746

RESULT 21
AD131610
ID AD131610 standard; cDNA; 2092 BP.
XX
XX AD131610;
XX
XX 17-JUN-2004 (first entry)
XX
XX Human cDNA #936.
XX
XX Human; gene; ss; immunological response; immunopathological condition;
KW Crohn's disease; asthma; ulcerative colitis; hyperesoinophilia;
KW irritable bowel syndrome; osteoarthritis; rheumatoid arthritis;
KW acute monocytic leukaemia; antinflammatory; antiasthmatic; antiulcer;
KW osteopathic; antiarthritic; antirheumatic; cytostatic.
XX
XX Homo sapiens.
XX
XX US6607879-B1.
XX
XX 19-AUG-2003.
XX
XX 09-FEB-1998; 98US-00023655.
XX
XX 09-FEB-1998; 98US-00023655.
XX
XX (INCY-) INCYTE CORP.
XX
XX Cocks BG, Stuart SG, Seilhamer JJ;
PI
```

```
XX WPI; 2003-895307/82.
XX
XX A composition comprising a plurality of cDNAs, useful for detecting
PT altered expression of genes in an immunological response or for
PT diagnosing and treating an immunopathology, e.g. Crohn's disease, asthma
PT or osteoarthritis.
XX
XX Claim 1; SEQ ID NO 936; 50pp; English.
XX
XX The invention relates to a composition comprising a plurality of cDNAs
CC for detecting the altered expression of genes in an immunological
CC response. The invention also relates to a method of diagnosing or
CC monitoring the treatment of an immunopathological condition in a sample,
CC comprising obtaining nucleic acids from a sample, contacting the nucleic
CC acids of the sample with an array comprising the plurality of cDNAs under
CC conditions to form one or more hybridisation complexes, detecting the
CC hybridisation complexes and comparing the levels of the detected
CC hybridisation complexes with the level of hybridisation complexes
CC detected in a non-diseased sample, where an altered level of the detected
CC hybridisation complexes correlates with the presence of an
CC immunopathological condition. Also disclosed are an expression profile
CC comprising a microarray and a plurality of detectable complexes and a
CC method for identifying a plurality of polynucleotide probes. The cDNAs
CC are useful as hybridisable array elements in a microarray for monitoring
CC the expression of target polynucleotides. The microarray can be used in
CC the diagnosis of an immunopathology, such as Crohn's disease, asthma,
CC ulcerative colitis, hyperesoinophilia, irritable bowel syndrome,
CC osteoarthritis, rheumatoid arthritis or acute monocytic leukaemia, and in
CC identifying agents for the treatment of the diseases. The microarray may
CC also be used in drug discovery and development, toxicological and
CC carcinogenicity studies, forensics or pharmacogenomics. The composition
CC may also be used in purification of a subpopulation of mRNAs, cDNAs or
CC genomic fragments. This sequence represents a human cDNA of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification but was obtained in electronic format directly
CC from USPTO at seqdata.uspto.gov/sequence.html.
XX
XX Sequence 2092 BP; 728 A; 359 C; 443 G; 562 T; 0 U; 0 Other;
SQ
Query Match 72.8%; Score 18.2; DB 11; Length 2092;
Best Local Similarity 87.0%; Pred. No. 3.3e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAACAACCTTGGGAAGGAGACT 23
Db 724 AAACAACACTTGGCAAGGAGACT 746

RESULT 22
AD583677
ID AD583677 standard; cDNA; 2092 BP.
XX
XX AD583677;
XX
XX 11-AUG-2005 (first entry)
XX
XX Human lymph node cDNA #936.
XX
XX ss; gene; human; immunological response; blood cell; cancer;
KW immunopathological; AIDS; allergy; anaemia; asthma; atherosclerosis;
KW bronchitis; ulcerative colitis; diabetes; multiple sclerosis;
KW osteoporosis; pancreatitis; infection; arthritis; lymph node.
XX
XX Homo sapiens.
XX
XX US2004077003-A1.
XX
XX 22-APR-2004.
XX
XX 14-AUG-2003; 2003US-00641643.
XX
XX 09-FEB-1998; 98US-00023655.
XX
XX
```

XX PA (INCY-) INCYTE CORP.
 XX PI Cocks BG, Stuart SG, Seilhamer JJ;
 XX DR WPI; 2004-387937/36.
 XX PT New compositions having a number of first, second and third
 XX PT polynucleotide probes, useful in research and diagnostic applications in
 XX PT cancer and immunological conditions e.g. AIDS, diabetes, osteoporosis and
 XX PT infections.
 XX PS Claim 15; SEQ ID NO 936; 16pp; English.
 XX CC The invention relates to polynucleotides which are used as probes to
 CC detect genes differentially expressed in an immunological response,
 CC abundantly expressed in an immunological response and/or coding for a
 CC polypeptide known to regulate blood cell biology. The polynucleotides are
 CC useful in research and diagnostic applications particularly in cancer and
 CC immunopathological conditions, such as AIDS, allergies, anaemia, asthma,
 CC atherosclerosis, bronchitis, ulcerative colitis, diabetes, multiple
 CC sclerosis, osteoporosis, pancreatitis, infections and arthritis. The
 CC present sequence represents a human lymph node cDNA used to detect blood
 CC cell and immunological response gene expression. Note: The present
 CC sequence does not appear in the printed specification but was obtained in
 CC electronic format from the USPTO web site
 CC (seqdata.uspto.gov/sequence.html?DocID=20040077003).
 XX SQ Sequence 2092 BP; 728 A; 359 C; 443 G; 562 T; 0 U; 0 Other;
 Query Match 72.8%; Score 18.2; DB 13; Length 2092;
 Best Local Similarity 87.0%; Pred. No. 3.3e+02;
 Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 AAAAAACACTTGGCAAGGAGACT 23
 |||||
 Db 724 AAACACACTTGGCAAGGAGACT 746

RESULT 23
 ADT47914/c
 ID ID ADT47914 standard; cDNA; 2306 BP.
 XX AC ADT47914;
 XX DT 02-DEC-2004 (first entry)
 XX DE Bacterial polynucleotide #22665.
 XX KW Recombinant DNA construct; transformed plant; improved plant property;
 KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
 KW pathogen tolerance; pest tolerance; plant disease resistance;
 KW cell cycle pathway modification; plant growth regulator;
 KW homologous recombination; seed oil yield; protein yield; carbohydrate;
 KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
 KW bacterial polynucleotide; gene; ss.
 XX OS Bacteria.
 XX PN US2003233675-A1.
 XX PD 18-DEC-2003.
 XX PF 20-FEB-2003; 2003US-00369493.
 XX PR 21-FEB-2002; 2002US-0360039P.
 XX PA (CAOY/) CAO Y.
 PA (HINK/) HINKLE G J.
 PA (SLAT/) SLATER S C.
 PA (CHEN/) CHEN X.
 PA (GOLD/) GOLDMAN B S.

PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
 XX WPI; 2004-061375/06.
 XX PT New recombinant DNA construct comprising a promoter positioned to provide
 PT for expression of a polynucleotide encoding a polypeptide from a
 PT microbial source, useful for producing plants with improved properties.
 XX Claim 1; SEQ ID NO 46352; 122pp; English.
 XX CC The invention relates to a recombinant DNA construct comprising a
 CC promoter functional in a plant cell, where the promoter is positioned to
 CC provide for expression of a polynucleotide encoding a polypeptide from a
 CC microbial source. The invention also relates to a transformed plant
 CC comprising the recombinant DNA construct and a method of producing a
 CC transformed plant having an improved property. The plant is a crop plant
 CC such as maize or soybean. The method of producing a transformed plant
 CC having an improved property comprises transforming a plant with the
 CC recombinant DNA construct and growing the transformed plant, where the
 CC polynucleotide or polypeptide is useful for improving plant properties.
 CC The recombinant DNA construct is useful for producing plants with
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,
 CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
 CC increased resistance to plant disease, better growth rate by modification
 CC of the cell cycle pathway with plant growth regulators, increased rate of
 CC homologous recombination, modified seed oil or protein yield and/or
 CC content, improved yield by modification of carbohydrate, nitrogen or
 CC phosphorus use and/or uptake, by modification of photosynthesis or by
 CC providing improved plant growth and development under at least one stress
 CC condition, improved lignin production or improved galactomannan
 CC production. This sequence represents a bacterial polynucleotide used in
 CC the scope of the invention. Note: The sequence data for this patent did
 CC not form part of the printed specification but was obtained in electronic
 CC format from USPTO at seqdata.uspto.gov/sequence.html.
 XX SQ Sequence 2306 BP; 689 A; 450 C; 447 G; 720 T; 0 U; 0 Other;
 Query Match 72.8%; Score 18.2; DB 13; Length 2306;
 Best Local Similarity 87.0%; Pred. No. 3.4e+02;
 Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 AAAAAACACTTGGGAAGGAGACT 24
 |||||
 Db 1161 ACAACACTAGGAAGGAGGCTC 1139

RESULT 24
 ACA57281
 ID ACA57281 standard; cDNA; 2307 BP.
 XX AC ACA57281;
 XX DT 10-JUN-2003 (first entry)
 XX DE Human adipocyte Selected Interacting domain, SID, cDNA #368.
 XX KW Human; ss; gene; prey; adipocyte; SID; selected interacting domain;
 KW anorectic; antidiabetic; protein-protein interaction; diabetes;
 KW yeast 2-hybrid assay; metabolic disorder; obesity.
 XX OS Homo sapiens.
 XX PN WO200286122-A2.
 XX PD 31-OCT-2002.
 XX PF 14-MAR-2002; 2002WO-EP003768.
 XX PR 14-MAR-2001; 2001US-0275734P.
 XX PA (HYBR-) HYBRIGENICS.
 XX PI Legrain P, Daviet L;

PT New nucleic acid differentially expressed in pancreatic tumor tissue, for
XX use as diagnostic agents and in screening for therapeutic agents.

PS Claim 1; SEQ ID NO 53; 28pp; German.

XX The invention relates to a novel human nucleic acid sequence of the
CC pancreas and its encoded protein. The invention further comprises:
CC proteins and peptides, preferably isolated, that contain a sequence
CC encoded by the novel nucleic acid; and methods for diagnosis and
CC treatment of pancreatic cancer; using a substance that inhibits or binds
CC to the protein or its DNA, including: an antisense oligonucleotide, short
CC interfering RNA or ribozyme directed against the pancreatic protein, an
CC organic molecule, particularly having a molecular weight below 5000,
CC especially 300, that binds to the pancreatic DNA, an aptamer or
CC (monoclonal) antibody, preferably human or humanized, that binds to the
CC pancreatic DNA, or an anti-idiotypic antibody raised against the
CC monoclonal antibody, any of which may be derivatized with a reporter
CC group, cytotoxic compound, immunostimulant and/or radioisotope. The novel
CC human pancreatic proteins and their encoding DNA have cytostatic
CC activity. The novel sequences are useful for inhibiting transcription
CC and/or expression of genes and proteins associated with pancreatic
CC cancer. This polynucleotide sequence represents the DNA encoding one of
CC the novel human pancreatic proteins of the invention. Note: This sequence
CC is not shown in the specification, it has been electronically downloaded
CC from a DVD-ROM provided with this specification by the European Patent
CC Office.

XX SQ Sequence 6721 BP; 2119 A; 1160 C; 1359 G; 2083 T; 0 U; 0 Other;

Query Match 72.8%; Score 18.2; DB 13; Length 6721;
Best Local Similarity 87.0%; Pred. No. 4e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AAAAAACACTTGGAGGAGACT 23

Db 1618 AAACACACTTGGCAAGGAGACT 1640

RESULT 27

ADZ49257

ID ADZ49257 standard; DNA; 6787 BP.

XX AC ADZ49257;

XX AC ADZ49257;

XX 30-JUN-2005 (first entry)

XX Insulin signaling pathway related gene, SEQ ID 586.

XX Antidiabetic; Gene Therapy; Non-insulin dependent diabetes;
XX insulin resistance; Gene; ds.

XX Homo sapiens.

XX US2005085436-A1.

XX 21-APR-2005.

XX 08-JUL-2004; 2004US-00887553.

XX 08-JUL-2003; 2003US-0485883P.

XX (LIHH//) LI H.

XX (MAJJ//) MA J.

XX Li H, Ma J;

XX WPI; 2005-305194/31.

XX Treating, preventing or ameliorating pathological conditions associated
XX with dysregulation of the insulin signaling pathway (ISP) comprises
XX administering to a subject an amount of a modulator of any of the
XX proteins regulated by ISP.

XX

PS Disclosure; SEQ ID NO 586; 70pp; English.

XX The present invention relates to a method for treating, preventing or
CC ameliorating pathological conditions associated with dysregulation of the
CC insulin signaling pathway (ISP). The method comprises administering to a
CC subject a modulator for ISP-regulated proteins or a pharmaceutical
CC composition comprising the described modulator. The method is useful for
CC treating, preventing or ameliorating pathological conditions associated
CC with dysregulation of the ISP such as Type II diabetes or Type A syndrome
CC of insulin resistance. The present sequence is a human homolog of a
CC Drosophila gene regulated by ISP. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20050085436.

XX SQ Sequence 6787 BP; 2134 A; 1172 C; 1388 G; 2093 T; 0 U; 0 Other;

Query Match 72.8%; Score 18.2; DB 14; Length 6787;
Best Local Similarity 87.0%; Pred. No. 4e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AAAAAACACTTGGAGGAGACT 23

Db 1685 AAACACACTTGGCAAGGAGACT 1707

RESULT 28

ADQ97164

ID ADQ97164 standard; DNA; 145985 BP.

XX AC ADQ97164;

XX 07-OCT-2004 (first entry)

XX Human cancer associated sequence HD1-08-009, SEQ ID 140.

XX Cytostatic; Gene Therapy; cancer; leukemia; lymphoma; Human; ds.

XX Homo sapiens.

XX WO2004060304-A2.

XX 22-JUL-2004.

XX 22-DEC-2003; 2003WO-US041389.

XX 27-DEC-2002; 2002US-00330773.

XX (SAGR-) SAGRES DISCOVERY INC.

XX Morris DW, Malandro MS;

XX WPI; 2004-543781/52.

XX New isolated cancer associated nucleic acids comprising at least 10
XX contiguous nucleotides, useful for diagnosing, preventing and/or treating
XX cancers such as leukemia and lymphoma.

XX Claim 1; SEQ ID NO 140; 199pp; English.

XX The present invention relates to cancer associated sequences (ADQ97025-
CC ADQ98004). The sequences are useful for the diagnosis, prevention and/or
CC treatment of cancer, such as leukemia and lymphoma. Note: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 145985 BP; 40325 A; 30740 C; 31191 G; 43391 T; 0 U; 338 Other;

Query Match 72.8%; Score 18.2; DB 12; Length 145985;
Best Local Similarity 87.0%; Pred. No. 6.1e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAAAGCTTGGAGGAGACT 23
 |||||
 DB 112891 AAAAAAAGCTTGGAGGAGATT 112913

RESULT 29
 ADD17061
 ID ADD17061 standard; DNA; 346 BP.
 XX
 AC ADD17061;
 XX
 DT 15-JAN-2004 (first entry)
 XX
 DE DNA (SeqID 1129) that confers an altered visual phenotype in plants.
 XX
 KW ds; visual phenotype; plant; architecture; leaf surface; chlorotic;
 KW bleaching; etching; wet leaf; stunting; elongation; texture;
 KW agronomic trait; growth regulation; dwarf variety; insect resistance;
 KW heat stress; transgenic.
 XX
 OS Unidentified.
 XX
 PN WO2003020741-A1.
 XX
 PD 13-MAR-2003.
 XX
 PF 30-AUG-2002; 2002WO-US027880.
 XX
 PR 31-AUG-2001; 2001US-0316326P.
 XX
 PA (DOWC) DOW CHEM CO.
 PA (DOWC) DOW AGROSCIENCES LLC.
 XX
 PI Crosley R, Skokut T, Ruegger M, Larrinua I, Shukla V;
 XX
 DR WPI; 2003-300858/29.
 XX
 PT Novel isolated nucleic acid derived from Nicotiana benthamiana, Oryzae
 PT sativa, Saccharomyces cerevisiae, and Papaver rhoeas, useful for
 PT conferring altered visual phenotypes in plants.
 XX
 PS Claim 1; SEQ ID NO 1129; 517pp; English.
 XX
 CC This invention relates to the identification and isolation of novel
 CC nucleic acid molecules that confer altered visual phenotypes in plants.
 CC Specifically, it refers to modifications of plant architecture and/ or
 CC leaf surface features in plants, such as chlorotic, bleaching, etching,
 CC wet leaf, stunting, elongation and texture phenotypes, which are thought
 CC will be agronomic traits beneficial to the farmer. As such, these novel
 CC phenotypes can affect growth regulation i.e. useful for creating dwarf
 CC varieties, exhibit resistance to insects or heat stress, confer changes
 CC in pigment content such that plants have enhanced vitamin production or
 CC delayed senescence and also for example produce plants that control the
 CC production of ethylene. Furthermore, the present invention comprises
 CC generating transgenic plants, as well as reproducibly altering the visual
 CC phenotype of plant seeds, plant tissues and plant cells containing the
 CC polynucleotides described herein. This polynucleotide is a homologue of a
 CC DNA sequence that confers an altered visual phenotype when expressed in
 CC plants, the method of the invention.
 XX
 SQ Sequence 346 BP; 111 A; 57 C; 84 G; 94 T; 0 U; 0 Other;

Query Match 71.2%; Score 17.8; DB 10; Length 346;
 Best Local Similarity 90.5%; Pred. No. 3.9e+02;
 Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 AACACTTGGAGGAGACTCA 25
 |||||
 DB 105 AACACTTGGATGAGACTCA 125

RESULT 30
 ADC76449

ID ADC76449 standard; DNA; 524 BP.
 XX
 AC ADC76449;
 XX
 DT 01-JAN-2004 (first entry)
 XX
 DE DNA homologous to phytopathogen resistance-related cDNA - SEQ ID 1718.
 XX
 KW rice; yeast; poppy; plant; disease resistance; anti-fungal;
 KW phytopathogen; gene shuffling; ds.
 XX
 OS Unidentified.
 XX
 PN WO2003020905-A2.
 XX
 PD 13-MAR-2003.
 XX
 PF 30-AUG-2002; 2002WO-US027883.
 XX
 PR 31-AUG-2001; 2001US-0316392P.
 XX
 PA (DOWC) DOW CHEM CO.
 XX
 PI Shukla V, Butler H, Larrinua I, Reddy AS;
 XX
 DR WPI; 2003-290185/28.
 XX
 PT Novel isolated nucleic acid derived from Nicotiana benthamiana, Oryzae
 PT sativa, Saccharomyces cerevisiae, Trichoderma harzianum and Papaver
 PT rhoeas, useful for conferring disease resistance in plants.
 XX
 PS Claim 1; SEQ ID NO 1718; 617pp; English.
 XX
 CC The invention relates to a novel isolated nucleic acid derived from
 CC Nicotiana benthamiana, Oryza sativa (rice), Saccharomyces cerevisiae
 CC (yeast), Trichoderma harzianum (Hypocrea lixii) and Papaver rhoeas
 CC (poppy) and a sequence that hybridises to them under conditions of low
 CC stringency, where expression of the nucleic acid in a plant results in a
 CC disease resistance phenotype. The polynucleotides of the invention
 CC demonstrate anti-fungal activity and may be useful in conferring disease
 CC resistance in a plant against phytopathogen such as Aspergillus flavus,
 CC Gibberella fujikuroi and Gibberella zeae. Furthermore, the
 CC polynucleotides may be useful to retrieve unknown sequences and in gene
 CC shuffling or sexual PCR procedures. The current sequence is that of the
 CC DNA of the invention which is homologous to that of the phytopathogen
 CC resistance-related contig cDNAs.
 XX
 SQ Sequence 524 BP; 169 A; 69 C; 138 G; 148 T; 0 U; 0 Other;

Query Match 71.2%; Score 17.8; DB 10; Length 524;
 Best Local Similarity 90.5%; Pred. No. 4.2e+02;
 Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 AACACTTGGAGGAGACTCA 25
 |||||
 DB 283 AACACTTGGATGAGACTCA 303

RESULT 31
 ADD17782
 ID ADD17782 standard; DNA; 524 BP.
 XX
 AC ADD17782;
 XX
 DT 15-JAN-2004 (first entry)
 XX
 DE DNA (SeqID 1850) that confers an altered visual phenotype in plants.
 XX
 KW ds; visual phenotype; plant; architecture; leaf surface; chlorotic;
 KW bleaching; etching; wet leaf; stunting; elongation; texture;
 KW agronomic trait; growth regulation; dwarf variety; insect resistance;
 KW heat stress; transgenic.
 XX

DR WPI; 2003-290185/28.
XX Novel isolated nucleic acid derived from *Nicotiana benthamiana*, *Oryzae*
PT *sativa*, *Saccharomyces cerevisiae*, *Trichoderma harzianum* and *Papaver*
PT *rheas*, useful for conferring disease resistance in plants.
XX
PS Claim 1; SEQ ID NO 774; 617pp; English.
XX
CC The invention relates to a novel isolated nucleic acid derived from
CC *Nicotiana benthamiana*, *Oryza sativa* (rice), *Saccharomyces cerevisiae*
CC (yeast), *Trichoderma harzianum* (*Hypocrea lixii*) and *Papaver rhoeas*
CC (poppy) and a sequence that hybridises to them under conditions of low
CC stringency, where expression of the nucleic acid in a plant results in a
CC disease resistance phenotype. The polynucleotides of the invention
CC demonstrate anti-fungal activity and may be useful in conferring disease
CC resistance in a plant against phytopathogens such as *Aspergillus flavus*,
CC *Gibberella fujikuroi* and *Gibberella zeae*. Furthermore, the
CC polynucleotides may be useful to retrieve unknown sequences and in gene
CC shuffling or sexual PCR procedures. The current sequence is that of the
CC DNA of the invention which is homologous to that of the phytopathogen
CC resistance-related contig cDNAs.
XX
SQ Sequence 530 BP; 173 A; 69 C; 140 G; 148 T; 0 U; 0 Other;

Query Match 71.2%; Score 17.8; DB 10; Length 530;
Best Local Similarity 90.5%; Pred. No. 4.2e+02; Indels 0; Gaps 0;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 AACACTTGGAGGGAGACTCA 25
DB 289 AACACTTGGATGAGACTCA 309
|||||

RESULT 34
ADB75375/c
ID ADB75375 standard; cDNA; 5347 BP.
XX
AC ADB75375;
XX
DT 04-DEC-2003 (first entry)
XX
DE Prostate cancer marker cDNA.
XX
KW Prostate; cancer; cytostatic; gene therapy; marker; ss.
XX
OS *Homo sapiens*.
XX
PN WO2003009814-A2.
XX
PD 06-FEB-2003.
XX
PF 25-JUL-2002; 2002WO-US023913.
XX
PR 25-JUL-2001; 2001US-0307982P.
PR 22-AUG-2001; 2001US-0314356P.
PR 25-SEP-2001; 2001US-0325020P.
PR 12-DEC-2001; 2001US-0341746P.
PR 05-MAR-2002; 2002US-0362158P.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Schlegel R, Monahan JE, Endege WO, Gannavarapu M, Gorbacheva B;
PI Hoerash S, Kamatkar S, Wonsay AM, Glatt K, Zhao X, Anderson D;
XX
PS WPI; 2003-248033/24.
XX
XX New nucleic acid molecule, useful for diagnosing or treating prostate
PT cancer.
XX
XX Disclosure; SEQ ID NO 199; 99pp; English.
PS
XX The invention relates to newly discovered cancer markers associated with
CC the cancerous state of prostate cells. Also disclosed is a method of

CC assessing whether a patient is afflicted with prostate cancer. The method
CC of the invention involves assessing whether a patient is afflicted with
CC prostate cancer by comparing the level of expression of a marker in a
CC patient sample and the normal level of expression of the marker in a
CC control non-prostate cancer sample, where a significant increase in the
CC level of expression of the marker in the patient sample and the normal
CC level indicates that the patient is afflicted with prostate cancer.
CC Nucleic acids of the invention are useful for diagnosing or treating
CC prostate cancer, and may be useful in gene therapy. Sequences given in
CC ADB75177-ADB75631 represent marker cDNA and proteins. Note: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 5347 BP; 1268 A; 1439 C; 1368 G; 1272 T; 0 U; 0 Other;

Query Match 71.2%; Score 17.8; DB 10; Length 5347;
Best Local Similarity 90.5%; Pred. No. 5.8e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAACAACCTTGGAGGGGAGA 21
DB 1038 AAGAAACACATGGAAGGGAGA 1018
|||||

RESULT 35
ABD33003/c
ID ABD33003 standard; DNA; 30642 BP.
XX
AC ABD33003;
XX
DT 18-NOV-2004 (first entry)
XX
DE Mouse cancer-associated genomic DNA MD21-026.
XX
KW Mouse; ds; cancer-associated protein; gene; cytostatic; cancer;
KW Leukaemia; lymphoma; CAP.
XX
OS *Mus musculus*.
XX
PN WO2004074320-A2.
XX
PD 02-SEP-2004.
XX
PF 17-FEB-2004; 2004WO-US004730.
XX
PR 14-FEB-2003; 2003US-00367094.
PR 14-MAR-2003; 2003US-00388838.
PR 15-APR-2003; 2003US-00417375.
PR 13-JUN-2003; 2003US-00461862.
PR 15-SEP-2003; 2003US-00663431.
PR 15-DEC-2003; 2003US-00737318.
XX
PA (SAGE-) SAGES DISCOVERY INC.
XX
PI Morris DW, Morris DW, Malandro MS;
XX
PS WPI; 2004-652914/63.
XX
XX New isolated cancer-associated polynucleotides and polypeptides useful
PT for diagnosing, preventing or treating cancers, especially lymphoma and
PT leukemia, or in screening for agents that modulate cancer.
XX
PS disclosure; seqid 773; 310pp; English.
XX
XX The invention relates to an isolated nucleic acid comprising at least 10
CC contiguous nucleotides of any of the 233 polynucleotide sequences given
CC in the specification, or its complement. The nucleic acids encode cancer-
CC associated proteins. Also included are an expression vector comprising
CC the isolated nucleic acid cited above, a host cell comprising the above
CC recombinant nucleic acid or expression vector, a microarray for detecting
CC a cancer-associated (CA) nucleic acid comprising at least one probe
CC comprising at least 10 contiguous nucleotides of any of the above-

CC mentioned nucleotide sequences, an isolated polypeptide (encoded within
CC an open reading frame of a CA sequence selected from any of the 95
CC polynucleotide sequences as mentioned in the specification, or its
CC complement), an isolated antibody, (or its antigen binding fragment) that
CC binds to the above polypeptide, a hybridoma that produces the above
CC monoclonal antibody, a pharmaceutical composition comprising the above
CC antibody and a pharmaceutical excipient, a kit for detecting cancer
CC cells (comprising the antibody cited above, methods for diagnosing cancer
CC or for detecting the presence or absence of cancer cells in an
CC individual, a method for inhibiting growth of cancer cells in an
CC individual, a method for delivering a therapeutic agent to cancer cells
CC in an individual, an electronic library comprising the above
CC polynucleotide or polypeptide (or their fragments), methods of screening
CC for anticancer activity or for a bioactive agent capable of modulating
CC the activity of a CA protein (CAP), methods for detecting cancer
CC associated with expression of a polypeptide in a test cell sample, a
CC method for treating cancers and a method for inhibiting the expression of
CC CA gene in a cell. The composition and methods are useful for detecting,
CC diagnosing, preventing and treating cancers, especially lymphoma and
CC leukemia. These may also be used in screening for agents that modulate
CC cancer. The present sequence is a mouse CAP genomic sequence. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences

SQ Sequence 30642 BP; 7617 A; 6784 C; 6669 G; 9572 T; 0 U; 0 Other;
Query Match 71.2%; Score 17.8; DB 13; Length 30642;
Best Local Similarity 90.5%; Pred. No. 7.5e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 AACACTTGGAGGGAGACTCA 25
Db 21362 AACACTTGGAGGGAGAGACA 21342

RESULT 36

ADQ97067/C
ID ADQ97067 standard; DNA; 72352 BP.

XX AC ADQ97067;
XX
XX
XX 07-OCT-2004 (first entry)
XX Mouse cancer associated sequence MD11-040, SEQ ID 43.
XX
XX Cytostatic; Gene Therapy; cancer; leukemia; lymphoma; Mouse; ds.

XX Mus musculus.

XX WO2004060304-A2.

XX 22-JUL-2004.

XX 22-DEC-2003; 2003WO-US0411389.

XX 27-DEC-2002; 2002US-00330773.

XX (SAGR-) SAGRES DISCOVERY INC.

XX Morris DW, Malandro MS;

XX WPI; 2004-543781/52.

XX New isolated cancer associated nucleic acids comprising at least 10
XX contiguous nucleotides, useful for diagnosing, preventing and/or treating
XX cancers such as leukemia and lymphoma.

XX Claim 1; SEQ ID NO 43; 199pp; English.

XX The present invention relates to cancer associated sequences (ADQ97025-
XX ADQ98004). The sequences are useful for the diagnosis, prevention and/or
XX treatment of cancer, such as leukemia and lymphoma. Note: The sequence

CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 72352 BP; 19658 A; 15474 C; 15337 G; 19982 T; 0 U; 1901 Other;

SQ Query Match 71.2%; Score 17.8; DB 12; Length 72352;
Best Local Similarity 90.5%; Pred. No. 8.5e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAACAACCTTGGAGGGAGA 21

Db 10509 AAAAACAACCTTGGAGGGAGA 10489

RESULT 37

ADQ36511
ID ADQ36511 standard; DNA; 169998 BP.

XX AC ADQ36511;

XX 09-AUG-2002 (first entry)

XX Human Her-1 gene.

XX Human; epidermal growth factor receptor; hyperproliferative disease;
XX Her1; psoriasis; psoriasis; tumour; cancer; gene; ds.

XX Homo sapiens.

XX Key Location/Qualifiers

FT exon 1208..1472

FT /tag= a

FT intron 1473..124390

FT /tag= b

FT exon 124391..124544

FT /tag= c

FT intron 124545..125409

FT /tag= d

FT exon 125410..125595

FT /tag= e

FT intron 125596..128711

FT /tag= f

FT exon 128712..128848

FT /tag= g

FT intron 128849..133400

FT /tag= h

FT exon 133401..133469

FT /tag= i

FT intron 133470..134652

FT /tag= j

FT exon 134653..134773

FT /tag= k

FT intron 134774..136116

FT /tag= l

FT exon 136117..136261

FT /tag= m

FT intron 136262..137936

FT /tag= n

FT exon 137937..138053

FT /tag= o

FT intron 138054..138637

FT /tag= p

FT exon 138638..138766

FT /tag= q

FT intron 138767..138864

FT /tag= r

FT exon 138865..138940

FT /tag= s

FT intron 138941..139765

FT /tag= t

FT exon 139766..139860

FT /tag= u

FT intron 139861. .142245
FT /tag= v
FT 142246. .142445
FT /tag= w
FT intron 142446. .143605
FT /tag= x
FT exon 143606. .143738
FT /tag= y
FT intron 143739. .145838
FT /tag= z
FT exon 145839. .145931
FT /tag= aa
FT intron 145932. .147385
FT /tag= ab
FT exon 147386. .147544
FT /tag= ac
FT intron 147545. .153274
FT /tag= ad
FT exon 153275. .153321
FT /tag= ae
FT intron 153322. .155088
FT /tag= af
FT exon 155089. .155231
FT /tag= ag
FT intron 155232. .156025
FT /tag= ah
FT exon 156026. .156151
FT /tag= ai
FT intron 156152. .156826
FT /tag= aj
FT exon 156827. .156928
FT /tag= ak
FT intron 156929. .163399
FT /tag= al
FT exon 163400. .163586
FT /tag= am

XX WO200226758-A1.

PN 04-APR-2002.

XX 28-SEP-2001; 2001WO-US030551.

XX 29-SEP-2000; 2000US-00676610.

XX (ISIS-) ISIS PHARM INC.

XX Bennett CF, Wyatt JR, Freier SM;

XX WPI; 2002-394234/42.

XX Novel antisense oligonucleotide that specifically hybridizes with and
XX inhibits nucleic acid encoding epidermal growth factor receptor, useful
XX for treating hyperproliferative disease such as cancer or psoriasis.

XX Example 19; Page 67-121; 169pp; English.

XX The invention relates to an antisense oligonucleotide targetted to a
XX nucleic acid molecule encoding human epidermal growth factor receptor
XX (Her1) to inhibit its expression. The antisense compounds are useful for
XX treating diseases or conditions associated with Her-1 such as
XX hyperproliferative diseases especially cancer (lung, ovarian, colon or
XX prostate cancer) and psoriasis. They are also useful as research
XX reagents, diagnostics, therapeutics, kits and prophylactically e.g. to
XX prevent or delay tumour formation. The present sequence is human Her-1
XX gene

XX Sequence 169998 BP; 46143 A; 38164 C; 37751 G; 47940 T; 0 U; 0 Other;

XX Query Match 71.2%; Score 17.8; DB 6; Length 169998;

XX Best Local Similarity 90.5%; Pred. No. 9.5e+02;

XX Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AAAAACACTTGGAGGGAGAC 22
Db 104105 AAAAGACTTGGAGGGAGAC 104125

RESULT 38

ABN85584

ID ABN85584 standard; DNA; 197496 BP.

XX AC ABN85584;

XX DT 09-SEP-2002 (first entry)

XX Human EGFR SEQ ID NO 10.

XX Human; EGFR; HER2-neu; chemotherapeutic regimen; tumour; cancer;
KW receptor tyrosine kinase; epidermal growth factor receptor;
KW gene expression; ds.

XX OS Homo sapiens.

XX PN WO200244413-A2.

XX PD 06-JUN-2002.

XX PF 09-NOV-2001; 2001WO-US043035.

XX PR 01-DEC-2000; 2000US-0250122P.

XX PR 04-DEC-2000; 2000US-0250469P.

XX PR 11-JUN-2001; 2001US-00877177.

XX PA (RESP-) RESPONSE GENETICS INC.

XX PI Danenberg KD;

XX DR WPI; 2002-537460/57.

XX Determining chemotherapeutic regimen of receptor tyrosine kinase targeted
XX agent for treating tumor by examining EGFR and/or HER2-neu mRNA amount in
XX tumor cells, comparing it to predetermined threshold expression level.

XX Disclosure; Page 71-124; 125pp; English.

XX The invention relates to determining the chemotherapeutic regimen of
XX receptor tyrosine kinase targeted agent for treating tumour by amplifying
XX mRNA from tumour and non-malignant tissues using a primer pair that
XX hybridises to epidermal growth factor receptor (EGFR) and/or HER2-neu
XX gene (i), quantitating and obtaining differential expression levels of
XX amplified mRNA and comparing the differential expression levels and
XX threshold levels for expression of (i). The method is useful for
XX assessment of clinical treatment of a patient and as a diagnostic or
XX prognostic tool for a range of cancers including breast, head and neck,
XX lung, oesophageal and colorectal cancer. The present sequence is that of
XX the human EGFR DNA sequence used in methods of the invention

XX Sequence 197496 BP; 53640 A; 44528 C; 43228 G; 56100 T; 0 U; 0 Other;

XX Query Match 71.2%; Score 17.8; DB 6; Length 197496;

XX Best Local Similarity 90.5%; Pred. No. 9.7e+02;

XX Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AAAAACACTTGGAGGGAGAC 22

Db 112105 AAAAGACTTGGAGGGAGAC 112125

RESULT 39

ABD32854

ID ABD32854 standard; DNA; 209083 BP.

XX AC ABD32854;

XX DT 18-NOV-2004 (first entry)

XX DE Human cancer-associated genomic DNA HD17-032.
XX KW Human; ds; cancer-associated protein; gene; cytostatic; cancer;
XX KW leukaemia; lymphoma; CAP.
XX OS Homo sapiens.
XX PN WO2004074320-A2.
XX PD 02-SEP-2004.
XX PF 17-FEB-2004; 2004WO-US004730.
XX PR 14-FEB-2003; 2003US-00367094.
XX PR 14-MAR-2003; 2003US-00388838.
XX PR 15-APR-2003; 2003US-00411735.
XX PR 13-JUN-2003; 2003US-00461862.
XX PR 15-SEP-2003; 2003US-00663431.
XX PR 15-DEC-2003; 2003US-00737318.
XX (SAGR-) SAGRES DISCOVERY INC.
XX PA Morris DW, Morris DW, Malandro MS;
XX PI WPI; 2004-652914/63.
XX DR New isolated cancer-associated polynucleotides and polypeptides useful
XX FT for diagnosing, preventing or treating cancers, especially lymphoma and
XX FT leukemia, or in screening for agents that modulate cancer.
XX PS claim 16; seqid 524; 310pp; English.
XX CC The invention relates to an isolated nucleic acid comprising at least 10
XX CC contiguous nucleotides of any of the 233 polynucleotide sequences given
XX CC in the specification, or its complement. The nucleic acids encode cancer-
XX CC associated proteins. Also included are an expression vector comprising
XX CC the isolated nucleic acid cited above, a host cell comprising the above
XX CC recombinant nucleic acid or expression vector, a microarray for detecting
XX CC a cancer-associated (CA) nucleic acid comprising at least one probe
XX CC comprising at least 10 contiguous nucleotides of any of the above-
XX CC mentioned nucleotide sequences, an isolated polypeptide (encoded within
XX CC an open reading frame of a CA sequence selected from any of the 95
XX CC polynucleotide sequences as mentioned in the specification, or its
XX CC complement), an isolated antibody, (or its antigen binding fragment) that
XX CC binds to the above polypeptide, a hybridoma that produces the above
XX CC monoclonal antibody, a pharmaceutical composition comprising the above
XX CC antibody and a pharmaceutical excipient, a kit for detecting cancer
XX CC cells (comprising the antibody cited above, methods for diagnosing cancer
XX CC or for detecting the presence or absence of cancer cells in an
XX CC individual, a method for inhibiting growth of cancer cells in an
XX CC individual, a method for delivering a therapeutic agent to cancer cells
XX CC in an individual, an electronic library comprising the above
XX CC polynucleotide or polypeptide (or their fragments), methods of screening
XX CC for anticancer activity or for a bioactive agent capable of modulating
XX CC the activity of a CA protein (CAP), methods for detecting cancer
XX CC associated with expression of a polypeptide in a test cell sample, a
XX CC method for treating cancers and a method for inhibiting the expression of
XX CC CA gene in a cell. The composition and methods are useful for detecting,
XX CC diagnosing, preventing and treating cancers, especially lymphoma and
XX CC leukaemia. These may also be used in screening for agents that modulate
XX CC cancer. The present sequence is a human CAP genomic sequence. Note: The
XX CC sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 209083 BP; 57124 A; 47057 C; 45679 G; 59223 T; 0 U; 0 Other;
Query Match 71.2%; Score 17.8; DB 13; Length 209083;
Best Local Similarity 90.5%; Pred. No. 9.7e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AAAAAGACTTGGAGGGAGAC 22

Db 112983 AAAAAGACTTGGAGGGAGAC 113003
RESULT 40
ADC24763/c
ID ADC24763 standard; DNA; 334462 BP.
XX AC ADC24763;
XX DT 18-DEC-2003 (first entry)
XX DE Human wild type HNL4X genomic sequence.
XX KW ds; nootropic; neuroleptic; tranquilizer; gene therapy; synaptogenesis;
XX KW mutation; neurological disease; mental disorder; psychiatric illness;
XX KW autism; Asperger syndrome; schizophrenia;
XX KW attention deficit hyperactivity disorder.
XX OS Homo sapiens.
XX PH Key Location/Qualifiers
XX FT exon 10670..10828
XX FT /tag= a
XX FT /number= 1b
XX FT intron 10829..83743
XX FT /tag= b
XX FT /number= 1b
XX FT exon 83744..84513
XX FT /tag= c
XX FT /number= 2
XX FT intron 84514..201675
XX FT /tag= d
XX FT /number= 2
XX FT exon 201676..201828
XX FT /tag= e
XX FT /number= 3
XX FT intron 201829..240358
XX FT /tag= f
XX FT /number= 3
XX FT exon 240359..240469
XX FT /tag= g
XX FT /number= 3b
XX FT intron 240470..312939
XX FT /tag= h
XX FT /number= 3b
XX FT exon 312940..313125
XX FT /tag= i
XX FT /number= 4
XX FT intron 313126..318312
XX FT /tag= j
XX FT /number= 4
XX FT exon 318313..319102
XX FT /tag= k
XX FT /number= 5
XX FT intron 319103..327017
XX FT /tag= l
XX FT /number= 5
XX FT exon 327018..330406
XX FT /tag= m
XX FT /number= 6
XX PN WO2003045998-A2.
XX PD 05-JUN-2003.
XX PF 02-DEC-2002; 2002WO-FR004134.
XX PR 30-NOV-2001; 2001CA-02364106.
XX PA (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
XX PA (INSP) INST PASTEUR.
XX PA (ASSI-) ASSISTANCE PUBLIQUE HOPITAUX PARIS.

XX Bourgeron T, Jamaïn S, Quach H, Betancur C, Leboyer M;
PI Gillberg C;
XX WPI; 2003-493399/46.
XX New nucleic acid encoding mutant protein involved in synaptogenesis,
PT useful for treatment and diagnosis of e.g. autism, Asperger syndrome, and
PT schizophrenia.
XX Claim 10; SEQ ID NO 1; 416pp; French.
XX The invention relates to an isolated or purified polynucleotide encoding
CC a polypeptide (the wild-type form of which is involved in synaptogenesis)
CC that includes at least one mutation associated with development of
CC neurological disease and/or a predisposition to development of mental
CC disorders or psychiatric illness. The polypeptide are used to screen for
CC agents that modulate their activity. Also nucleic acid, polypeptide,
CC polypeptide-specific antibodies, vectors containing he nucleic acid and
CC host cells containing the vector, are useful as pharmaceuticals for
CC treating mental and neurological disorders, specifically autism, Asperger
CC syndrome, schizophrenia and attention deficit hyperactivity disorder. The
CC wild-type forms of the nucleic acid and polypeptide can be used
CC similarly. Also detecting mutations in the nucleic acid and polypeptide,
CC or measuring activity of the polypeptide, can be used to detect
CC biochemical disorders that affect formation of synapses and to diagnose
CC mental disease. This sequence corresponds to the genomic sequence of the
CC human wild type HNL4X gene.
XX Sequence 334462 BP; 97253 A; 64494 C; 66298 G; 106417 T; 0 U; 0 Other;
SQ Query Match 71.2%; Score 17.8; DB 10; Length 334462;
Best Local Similarity 90.5%; Pred. No. 1e+03;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 AAAAAACACTTGGAGGGAGA 21
Db 198779 AATAACATTGGAGGGAGA 198759
RESULT 41
AAC13345/C
ID AAC13345 standard; cDNA; 120 BP.
XX AAC13345;
XX 06-OCT-2000 (first entry)
XX Human secreted protein 5' EST, SEQ ID NO: 17420.
XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping; ss.
XX Homo sapiens.
XX EP1033401-A2.
XX 06-SEP-2000.
XX 21-FEB-2000; 2000EP-00200610.
XX 26-FEB-1999; 99US-0122487P.
XX (GEST) GENSET.
XX Dumas Milne Edwards J, Duclert A, Giordano J;
PI WPI; 2000-500381/45.
XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures.
XX

PS Claim 1; SEQ ID NO 17420; 71pp + Sequence Listing; English.
XX The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC identified within the present sequence. The 5' ESTs were prepared from
CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
CC sequences usually correspond mainly to the 3' untranslated region (UTR)
CC of the mRNA because they are often obtained from oligo-dT primed cDNA
CC libraries. Such ESTs are not well suited for isolating cDNA sequences
CC derived from the 5' ends of mRNAs and even in those cases where longer
CC cDNA sequences have been obtained, the full 5' UTR is rarely included. 5'
CC ESTs are derived from mRNAs with intact 5' ends and can therefore be used
CC to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in
CC diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC They are used to obtain upstream regulatory sequences and to design
CC expression and secretion vectors
XX Sequence 120 BP; 29 A; 25 C; 23 G; 43 T; 0 U; 0 Other;
SQ Query Match 70.4%; Score 17.6; DB 3; Length 120;
Best Local Similarity 83.3%; Pred. No. 4.2e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 AAAAAACACTTGGAGGGAGACTC 24
Db 114 AAAAAAACTTGTAAAGTTAGACTC 91
RESULT 42
AAV04596
ID AAV04596 standard; DNA; 299 BP.
XX AAV04596;
XX 25-MAR-2003 (revised)
DT 02-JUL-1998 (first entry)
XX Flea serine protease DNA sequence SEQ ID NO:127.
XX Flea; Siphonaptera; protease; vaccine; prophylaxis; infestation;
KW immunoglobulin protease; larvae; host animal; ss.
XX Siphonaptera.
XX WO9740058-A1.
XX 30-OCT-1997.
XX 24-APR-1997; 97WO-US0006121.
XX 24-APR-1996; 96US-00639075.
PR 15-NOV-1996; 96US-00749699.
PR 04-APR-1997; 97US-0042945P.
XX (HESK-) HESKA CORP.
XX Grieve RB, Rushlow KE, Hunter SW, Frank GR, Steigler GL;
PI Gaines PJ, Silver G;
XX WPI; 1998-076762/07.
XX New flea protease genes and proteins - used in vaccine compositions for
PT the prophylaxis and treatment of flea infestation, especially in cats or
PT dogs.
XX Claim 1; Page 259; 318pp; English.
XX The present sequence is a DNA sequence from a novel flea serine protease.
CC The protease, its mimetopes, antibodies (Ab) and inhibitors of the
CC protein, as well as the DNA encoding the protein, may all be used in
CC therapeutic compositions to reduce flea protease activity (especially
CC immunoglobulin protease) and so reduce flea infestation, especially in
CC cats or dogs. Alternatively, flea larvae may ingest the faeces of adult
CC

CC fleas which comprises anti-protease antibodies produced by a host animal
 CC in response to administration of the protein. Therapeutic compositions
 CC may further comprise a compound that reduces haematophagous ectoparasite
 CC burden by a method other than by reducing flea immunoglobulin protease
 CC activity. The novel flea DNA encoding the protein can also be used to
 CC produce recombinant protein, and fragments of it are used as probes and
 CC primers for identification and isolation of related sequences, also as
 CC antisense, triplex-forming agents and ribozymes for inhibition of the
 CC synthesis of the protein. Ab are also useful for screening expression
 CC libraries, to purify the protein and to target cytotoxins to fleas.
 CC (Updated on 25-MAR-2003 to correct PI field.)

XX
 SQ Sequence 299 BP; 94 A; 45 C; 79 G; 79 T; 0 U; 2 Other;

Query Match 70.4%; Score 17.6; DB 2; Length 299;
 Best Local Similarity 83.3%; Pred. No. 4.7e+02;
 Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AAAACACTTGGGAAGGGAGACTCA 25
 |||||
 Db 267 AAAGACACTTGCAGGGAGATTCA 290

RESULT 43

AAAC90858
 ID AAC90858 standard; cDNA; 299 BP.

XX AAC90858;

XX 19-MAR-2001 (first entry)

DE Flea serine protease nFSP8-299 nucleotide sequence #92.

XX Flea; cat flea; serine protease; aminopeptidase; cysteine protease;
 KW flea infestation; proteolytic; insecticide; vaccine; cat; dog; dog flea;
 KW Ctenocephalides felis; Ctenocephalides canis; Pulex irritans;
 KW ectoparasite; ss.

XX Ctenocephalides felis.

XX US6150125-A.

XX 21-NOV-2000.

XX 24-APR-1996; 96US-00639075.

XX 13-DEC-1991; 91US-00806482.

XX 18-OCT-1994; 94US-00326773.

XX 07-JUN-1995; 95US-00482130.

XX 07-JUN-1995; 95US-00484211.

XX 07-JUN-1995; 95US-00485443.

XX 15-AUG-1997; 97WO-US014442.

XX (HESK-) HESKA CORP.

XX Gaines PV, Silver G, Rushlow KE, Hunter SW, Frank GR;

PI Stiegler GL, Grieve RB;

XX WPI; 2001-136374/14.

DR P-PSDB; AAB50602.

XX New isolated flea proteins with proteolytic activity, useful for
 PT preventing and reducing flea infestations in mammals especially cats and
 PT dogs.

XX Claim 4; Col 153-154; 150pp; English.

XX The present invention describes isolated flea serine protease,
 CC aminopeptidase and cysteine protease proteins (I). Also described is a
 CC method for identifying a compound (II) capable of inhibiting flea
 CC protease activity comprising: (a) contacting (I) with a protease
 CC substrate and a putative inhibitory compound, where (I) has proteolytic

CC activity in the absence of the compound; and (b) determining if the
 CC compound inhibits protease activity by detecting cleavage of the protease
 CC substrates; where decreased cleavage of the protease substrate indicates
 CC an inhibitory compound. (I), nucleic acid molecules encoding (I), and
 CC antibodies immunospecific for (I) and (II) are useful for preventing and
 CC reducing flea infestations, particularly the species Ctenocephalides
 CC felis (cat fleas), Ctenocephalides canis (dog fleas) and Pulex irritans,
 CC in animals, preferably cats and dogs. They are also useful for reducing
 CC infestation by other ectoparasites, preferably mosquitoes, midges,
 CC sandflies, blackflies, ticks and Rhodnius. AAC90818 to AAC90913 and
 CC AAB50551 to AAB50644 represent sequences used in the exemplification of
 CC the present invention

XX
 SQ Sequence 299 BP; 94 A; 45 C; 79 G; 79 T; 0 U; 2 Other;

Query Match 70.4%; Score 17.6; DB 4; Length 299;
 Best Local Similarity 83.3%; Pred. No. 4.7e+02;
 Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AAAACACTTGGGAAGGGAGACTCA 25
 |||||
 Db 267 AAAGACACTTGCAGGGAGATTCA 290

RESULT 44

AAA00406/c

ID AAA00406 standard; cDNA; 300 BP.

XX AAA00406;

XX 19-MAY-2000 (first entry)

XX Human colon cancer cell line polynucleotide sequence SEQ ID NO:397.

XX Human; colon cancer; tumour; diagnosis; gene expression product; probe;
 KW detection; cancerous state; metastasis; identification; breast cancer;
 KW oestrogen receptor-positive breast cancer; therapy;
 KW oestrogen receptor-negative breast cancer; lung cancer; ss.

XX Homo sapiens.

XX WO9558675-A2.

XX 18-NOV-1999.

XX 13-MAY-1999; 99WO-US010602.

XX 14-MAY-1998; 98US-0085426P.

XX 15-MAY-1998; 98US-0085537P.

XX 15-MAY-1998; 98US-0085696P.

XX 21-OCT-1998; 98US-0105234P.

XX 27-OCT-1998; 98US-0105877P.

XX (CHIR) CHIRON CORP.

XX (HYSE-) HYSEQ INC.

XX Williams LT, Escobedo J, Innis MA, Garcia PD, Sudduth-Klinger J;

PI Reinhard C, Giese K, Randazzo F, Kennedy GC, Pot D, Kassam A;

PI Lamson G, Drmanac R, Crkvenjakov R, Dickson M, Drmanac S, Labat I;

PI Leshkowitz D, Kita D, Garcia V, Jones LW, Stache-Crain B;

XX WPI; 2000-126369/11.

XX Polynucleotide library used to determine cancerous states of mammalian
 PT cells.

XX Claim 1; Page 275; 1097pp; English.

XX AAA00010 to AAA02716 represent polynucleotides isolated from cDNA
 CC libraries constructed from human colon cancer cell lines. The present
 CC invention also describes a method of detecting differentially expressed
 CC genes correlated with a cancerous state of a mammalian cell, comprising
 CC detecting at least one differentially expressed gene product in a test

CC sample derived from a cell suspected of being cancerous, where detection
CC of the differentially expressed gene product is correlated with a
CC cancerous state of the cell from which the test sample was derived. The
CC polynucleotide sequences can be used in a method for detecting
CC differentially expressed genes correlated with a cancerous state of a
CC mammalian cell. The polynucleotides can also be used as probes for
CC detecting and mapping related genes. They can be used in diagnosis and
CC prognosis of diseases and disorders (e.g. identification of pre-
CC metastatic or metastatic cancerous states, stages of cancer, or
CC responsiveness of cancer to therapy). This is particularly for breast
CC cancer, oestrogen receptor-positive breast cancer, oestrogen receptor-
CC negative breast cancer, lung cancer, and colon cancer
XX
SQ Sequence 300 BP; 88 A; 37 C; 47 G; 128 T; 0 U; 0 Other;

Query Match 70.4%; Score 17.6; DB 3; Length 300;
Best Local Similarity 83.3%; Pred. NO. 4.7e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAACACTTGGAGGGAGACTC 24
Db 294 AGAAAAACACTTGGAGGAAAGCTC 271

RESULT 45
AAF65920/c
ID AAF65920 standard; cDNA; 389 BP.
XX
AC AAF65920;
XX
XX 09-APR-2001 (first entry)
XX
DE Novel human polynucleotide, SEQ ID NO: 1676.
XX
KW Human; cytostatic; gene therapy; colon cancer; prostate cancer;
KW breast cancer; lung cancer; cancer detection; ss.
XX
OS Homo sapiens.
XX
XX WO200102568-A2.
PN
PD 11-JAN-2001.
XX
XX 30-JUN-2000; 2000WO-US018374.
PF
XX 02-JUL-1999; 99US-0142310P.
PR
XX 02-JUL-1999; 99US-0142311P.
XX
PA (CHIR) CHIRON CORP.
PA (HYSE-) HYSEQ INC.
XX
XX Williams LT, Escobedo J, Innis MA, Garcia PD, Klinger J;
PI Kassam A, Reinhard C, Randazzo F, Kennedy GC, Pot D, Lamson G;
PI Drmanac R, Crkenjakov R, Drmanac S, Dickson M, Labat I;
PI Leshkowitz D, Kita D, Garcia V, Jones LW, Strache-Crain B;
XX
XX WPI; 2001-091805/10.
DR
XX Library of polynucleotides for diagnosing a cancerous state of a
PT mammalian cell and detecting cancer, particularly of the colon or
PT prostate, comprises 3351 human polynucleotide sequences.
PT
XX Claim 9; Page 784; 1046pp; English.
PS
XX The present sequence is one of 3351 sequences in a library of human
CC polynucleotides. The library is used to detect differentially expressed
CC genes correlated with a cancerous state of a mammalian cell and can
CC detect colon, prostate, breast and lung cancer. The library can be used
CC to produce probes for detection of mRNA and to produce additional copies
CC of the polynucleotides. The probes can be used for chromosome mapping of
CC the polynucleotide and for detection of transcription levels. Ribozymes
CC or antisense oligonucleotides can be generated. The polynucleotides and
CC their gene products are used as genetic or biochemical markers (e.g. in

CC blood or tissues) that will detect the earliest changes along the
CC carcinogenesis pathway and/or monitor the efficacy of therapies and
CC preventive interventions. The polynucleotides, polypeptides and
CC antibodies against them can be used in pharmaceutical compositions to
CC treat the cancers and proliferative disorders such as neoplasia,
CC dysplasia and hyperplasia
XX
SQ Sequence 389 BP; 109 A; 87 C; 93 G; 99 T; 0 U; 1 Other;

Query Match 70.4%; Score 17.6; DB 5; Length 389;
Best Local Similarity 83.3%; Pred. NO. 4.9e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAACACTTGGAGGGAGACTC 24
Db 242 ATAAAGGACTTGGAAAGAGAGACTC 219

RESULT 46
AAI86689/c
ID AAI86689 standard; cDNA; 398 BP.
XX
AC AAI86689;
XX
XX 06-NOV-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 8749.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation; ss.
XX
OS Homo sapiens.
XX
XX WO200164835-A2.
PN
PD 07-SEP-2001.
XX
XX 26-FEB-2001; 2001WO-US004927.
PF
XX 28-FEB-2000; 2000US-00515126.
PR
XX 18-MAY-2000; 2000US-00577409.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Drmanac RT;
PI
XX WPI; 2001-514838/56.
DR
XX P-PSDB; AAO08758.
DR
XX Isolated nucleic acids and polypeptides, useful for preventing diagnosing
PT and treating e.g. leukemia, inflammation and immune disorders.
PT
XX Claim 1; SEQ ID NO 8749; 1399pp + Sequence Listing; English.
PS
XX The invention relates to human polynucleotides (AAI79941-AAI93841) and
CC the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 398 BP; 123 A; 69 C; 74 G; 131 T; 0 U; 1 Other;
XX
Query Match 70.4%; Score 17.6; DB 4; Length 398;

CC 2003 to correct PI field.)

XX Sequence 436 BP; 139 A; 64 C; 111 G; 118 T; 0 U; 4 Other;

SX

Query Match 70.4%; Score 17.6; DB 2; Length 436;

Best Local Similarity 83.3%; Pred. No. 5e+02;

Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps

QY 2 AAAAACACTTGAGAGGAGACTCA 25
||| ||||| ||||| |||

Db 254 AAAGACACTTGCAGGAGATTCA 277

RESULT 48

AAC90876

ID AAC90876 standard; cDNA; 436 BP.

XX AAC90876;

XX

DT 19-MAR-2001 (first entry)

XX

DE Flea serine protease nfsP8-436 nucleotide sequence #122.

XX

KW Flea; cat flea; serine protease; aminopeptidase; cysteine protease;

KW flea infestation; proteolytic; insecticide; vaccine; cat; dog; dog flea;

KW Ctenocephalides felis; Ctenocephalides canis; Pulex irritans;

KW ectoparasite; ss.

XX

OS Ctenocephalides felis.

XX

US6150125-A.

XX

XX 21-NOV-2000.

XX

XX 24-APR-1996; 96US-00639075.

XX

PR 13-DEC-1991; 91US-00806482.

PR 18-OCT-1994; 94US-00326773.

PR 07-JUN-1995; 95US-00482130.

PR 07-JUN-1995; 95US-00484211.

PR 07-JUN-1995; 95US-00485443.

PR 07-JUN-1995; 95US-00485455.

PR 15-AUG-1997; 97WO-US014442.

XX

PA (HESK-) HESKA CORP.

XX

PI Gaines PJ, Silver G, Rushlow KE, Hunter SW, Frank GR;

PI Stiegler GL, Grieve RB;

XX

WPI; 2001-136374/14.

DR

DR P-PSDB; AAB50614.

XX

XX New isolated flea proteins with proteolytic activity, useful for preventing and reducing flea infestations in mammals especially cats and dogs.

PT

PT

PS Claim 4; Col 189-190; 150pp; English.

XX

XX

XX The present invention describes isolated flea serine protease, aminopeptidase and cysteine protease proteins (I). Also described is a method for identifying a compound (II) capable of inhibiting flea protease activity comprising: (a) contacting (I) with a protease substrate and a putative inhibitory compound, where (I) has proteolytic activity in the absence of the compound; and (b) determining if the compound inhibits protease activity by detecting cleavage of the protease substrates; where decreased cleavage of the protease substrate indicates an inhibitory compound. (I), nucleic acid molecules encoding (I), and antibodies immunospecific for (I) and (II) are useful for preventing and reducing flea infestations, particularly the species Ctenocephalides felis (cat fleas), Ctenocephalides canis (dog fleas) and Pulex irritans, in animals, preferably cats and dogs. They are also useful for reducing infestation by other ectoparasites, preferably mosquitoes, midges, sandflies, blackflies, ticks and Rhodnius. AAC90818 to AAC90913 and

CC

CC AAB50551 to AAB50644 represent sequences used in the exemplification of
CC the present invention
XX
SQ Sequence 436 BP; 139 A; 64 C; 111 G; 118 T; 0 U; 4 Other;

Query Match 70.4%; Score 17.6; DB 4; Length 436;
Best Local Similarity 83.3%; Pred. No. 5e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AAAACACTTGGAGGGAGACTCA 25
DB 254 AAAGACACTTGCAGGGAGATTCA 277

RESULT 49
AABF83431
ID AAF83431 standard; cDNA; 436 BP.
XX
AC AAF83431;
XX
DT 09-JUL-2001 (first entry)
XX
DE Flea serine protease coding sequence nfp8_436.
XX
KW Serine protease; flea; hematophagous; ectoparasite; infestation; vaccine;
KW immune response; protease; ss.
XX
OS Ctenocephalides sp.
XX
FH Key Location/Qualifiers
FT CDS 2..4367
FT /*tag= a
XX
TX US6204010-B1.
PN 20-MAR-2001.
XX
XX 27-FEB-1998; 98US-00032215.
XX
PR 07-JUN-1995; 95US-00484211.
PR 18-OCT-1995; 95WO-US014442.
PR 24-APR-1996; 96US-00639075.
PR 15-NOV-1996; 96US-00749699.
PR 24-APR-1997; 97WO-US006121.
PR 01-AUG-1997; 97US-00817795.
PR 14-NOV-1997; 97US-00970995.
XX
PA (HESK-) HESKA CORP.
XX
PI Stiegler GL, Gaines PJ;
XX
XX WPI; 2001-289510/30.
DR P-PSDB; AAB62505.
XX
PT New flea serine proteases and nucleic acids encoding them, useful for
PT reducing hematophagous ectoparasite, particularly flea infestation, and
PT as vaccines for protecting an animal from flea infestation.
XX
XX Example 1; Col 63-64; 66pp; English.
XX
CC The invention relates to flea serine protease proteins and nucleic acid
CC molecules encoding the serine proteases. The nucleic acids are useful for
CC reducing hematophagous ectoparasite, particularly flea infestation, and
CC as vaccines for protecting the animal from flea infestation. The nucleic
CC acids are also useful for eliciting an immune response against a flea
CC serine protease protein. The flea protease proteins may be used to
CC identify additional protease inhibitors, which can be used to identify
CC preferred types of flea proteases, and to purify corresponding proteases.
CC The present sequence represents a flea serine protease nucleic acid
XX sequence nfp8_436
XX
SQ Sequence 436 BP; 139 A; 64 C; 111 G; 118 T; 0 U; 4 Other;

Query Match 70.4%; Score 17.6; DB 5; Length 436;
Best Local Similarity 83.3%; Pred. No. 5e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AAAACACTTGGAGGGAGACTCA 25
DB 254 AAAGACACTTGCAGGGAGATTCA 277

RESULT 50
ACL63156/c
ID ACL63156 standard; cDNA; 480 BP.
XX
AC ACL63156;
XX
DT 24-MAR-2005 (first entry)
XX
DE Human colon cancer differentially expressed polynucleotide, SEQ ID:9291.
XX
KW Differential expression; diagnosis; therapy; drug screening; cancer;
KW neoplasm; colon tumor; breast tumor; pancreas tumor; cytostatic; vaccine;
KW ss.
XX
OS Homo sapiens.
XX
PN WO2005000087-A2.
XX
PD 06-JAN-2005.
XX
XX 13-MAY-2004; 2004WO-US015421.
XX
PR 03-JUN-2003; 2003US-0475872P.
XX
PA (CHIR) CHIRON CORP.
XX
PI Randazzo F, Moler E, Escobedo J, Garcia PD;
XX
XX WPI; 2005-075421/08.
XX
PT New isolated polynucleotides, which are differentially expressed in colon
PT cancer cell, useful for treating cancer, e.g. colon cancer, breast
PT cancer, or pancreatic cancer.
XX
PS Claim 1; SEQ ID NO 9291; 97pp; English.
XX
CC The invention relates to 9672 polynucleotides (ACL53866-ACL63537) which
CC are differentially expressed in colon cancer cells. The invention also
CC relates to vectors and host cells comprising a differentially expressed
CC polynucleotide of the invention; a method for detecting a cancerous cell
CC by detection of a gene product of the polynucleotides; a method for
CC inhibiting a cancerous phenotype of a cell by inhibiting a gene product
CC of the polynucleotides; a method of treating an individual with cancer by
CC administration of a modulator of a gene product of the polynucleotides;
CC and an isolated antibody that specifically binds to a polypeptide encoded
CC by one of the 9672 polynucleotides. The polynucleotides, polypeptides,
CC antibodies, and methods are useful for the detection of cancerous cells;
CC for the diagnosis, prognosis and management of cancer; for the
CC identification of agents that modulate the phenotype of cancerous cells;
CC for the identification of therapeutic targets for cancer chemotherapy;
CC and for the treatment of cancer, especially colon cancer and metastasized
CC colon cancer, but also breast or pancreatic cancer. The polynucleotides
CC are also useful as a source of probes or primers for use in diagnostic
CC methods. The differentially expressed polynucleotides or their encoded
CC proteins can additionally be used as vaccines to modulate primary immune
CC responses for the prevention or treatment of cancer. The present sequence
CC represents a specifically claimed polynucleotide which is differentially
CC expressed in colon cancer. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 480 BP; 127 A; 123 C; 102 G; 128 T; 0 U; 0 Other;

Query Match 70.4%; Score 17.6; DB 14; Length 480;
Best Local Similarity 83.3%; Pred. No. 5.1e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 AAAAAACACTTGAAGGAGACTC 24
| | | | | | | | | | | | | | | | | | | | | |
Db 239 ATAAAGGACTTGAAGGAGACTC 216

Search completed: February 3, 2006, 21:56:41
Job time : 211.111 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 3, 2006, 14:38:25 ; Search time 2951.67 Seconds
(without alignments)
396.277 Million cell updates/sec

Title: US-10-719-900-10

Perfect score: 25

Sequence: 1 aaaaacacttggaaggagactca 25

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database :

EST:*

1: gb_est1.*

2: gb_est2.*

3: gb_est3.*

4: gb_est4.*

5: gb_est5.*

6: gb_est6.*

7: gb_est7.*

8: gb_est8.*

9: gb_est9.*

10: gb_est10.*

11: gb_est11.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	23.4	93.6	450	4	AK007298 Mus muscu
2	23.4	93.6	460	5	BY707400 BY707400
3	20.8	83.2	298	5	BX631074 BX631074
4	20.8	83.2	464	2	BF796159 BF796159
5	20.8	83.2	466	2	BF454696 BF454696
6	20.8	83.2	623	10	CL346540 CL346540
7	20.8	83.2	666	2	BF454299 BF454299
8	20.8	83.2	669	2	BF538629 BF538629
9	20.8	83.2	680	2	BF872391 BF872391
10	20.8	83.2	682	2	BF865082 BF865082
11	20.8	83.2	683	2	BF866198 BF866198
12	20.8	83.2	737	2	BF866547 BF866547
13	20.8	83.2	747	2	BF866078 BF866078
14	20.8	83.2	755	2	BF866256 BF866256
15	20.8	83.2	764	7	CK294820 CK294820
16	20.8	83.2	775	2	BF870364 BF870364
17	20.8	83.2	856	2	BF865927 BF865927
18	20.8	83.2	881	2	BF302440 BF302440
19	20.8	83.2	904	2	BF538107 BF538107
20	20.8	83.2	928	2	BF537576 BF537576
21	20.8	83.2	952	2	BF302117 BF302117
22	20.2	80.8	711	9	BZ023877 BZ023877

96	18.8	75.2	932	7	CK402743	AUF IfInt
97	18.8	75.2	963	8	DN577957	922f6236
98	18.8	75.2	983	9	CC331869	GUEN69TV
99	18.8	75.2	1187	8	DN701279	CLJ34-G01
100	18.8	75.2	1334	10	AG442056	Mus muscu
101	18.8	75.2	2829	4	AK081536	Mus muscu
102	18.6	74.4	141	7	CV307827	tj47b09.b
103	18.6	74.4	141	7	CV307828	tj47b09.g
104	18.6	74.4	336	7	CK713568	2F201-P00
105	18.6	74.4	351	8	CK555204	Yda97d08.
106	18.6	74.4	400	3	BI790314	IC96h07.x
107	18.6	74.4	403	8	DR903915	JGI_XZT65
108	18.6	74.4	416	8	DR542611	WS01039.B
109	18.6	74.4	440	6	CD313383	StrPu621.
110	18.6	74.4	446	1	AA403014	zC63612.f
111	18.6	74.4	452	5	BU671855	HaT423 He
112	18.6	74.4	459	1	AA398499	zt63612.s
113	18.6	74.4	508	5	BQ386941	NISC mn21
114	18.6	74.4	512	9	AZ098943	RPCI-23-4
115	18.6	74.4	516	9	AQ998152	RPCI-23-2
116	18.6	74.4	517	9	BZ330770	hV97a08.g
117	18.6	74.4	517	9	AQ685663	HS_5556.A
118	18.6	74.4	536	5	BX116611	BX116611
119	18.6	74.4	543	3	BM257260	520815.MA
120	18.6	74.4	545	6	CD332039	StrPu537.
121	18.6	74.4	558	1	AW441225	EST310621
122	18.6	74.4	577	9	CE130490	tigr-g88-
123	18.6	74.4	584	9	AZ640544	1M0502C09
124	18.6	74.4	593	9	BH281646	CH230-49G
125	18.6	74.4	609	9	AZ080614	RPCI-23-4
126	18.6	74.4	618	10	CZ094521	OM_Ba009
127	18.6	74.4	619	10	CE525050	tigr-g88-
128	18.6	74.4	621	8	CV950784	Pvrvb43
129	18.6	74.4	636	6	CD307379	StrPu691.
130	18.6	74.4	652	11	CR085510	Reverse s
131	18.6	74.4	654	10	CB413313	tigr-g88-
132	18.6	74.4	659	6	CB530282	737067.WA
133	18.6	74.4	663	10	CE537115	tigr-g88-
134	18.6	74.4	667	10	CE752213	tigr-g88-
135	18.6	74.4	673	7	CK950704	4089959.B
136	18.6	74.4	679	6	CD336366	StrPu537.
137	18.6	74.4	680	6	CD310869	StrPu691.
138	18.6	74.4	689	10	CE677158	tigr-g88-
139	18.6	74.4	709	9	CC390114	PuHT0127B
140	18.6	74.4	718	9	CE138835	tigr-g88-
141	18.6	74.4	744	8	DR439912	EST149.18
142	18.6	74.4	752	10	AG376010	Mus muscu
143	18.6	74.4	763	1	AJ814183	AJ814183
144	18.6	74.4	776	2	BG203562	RST22946
145	18.6	74.4	780	10	AG603339	Mus muscu
146	18.6	74.4	783	2	BG213532	RST33139
147	18.6	74.4	784	10	CL581889	OB_Ba004
148	18.6	74.4	799	2	BG189152	RST8192.A
149	18.6	74.4	838	10	DU039116	22937.Tom
150	18.6	74.4	865	9	BZ722675	PUDAL237B

ALIGNMENTS

RESULT 1
AK007298 450 bp mRNA linear HTC 03-APR-2004
LOCUS Mus musculus adult male testis cDNA, RIKEN full-length enriched
DEFINITION library, clone:1700127D06 product:similar to tissue kallikrein (EC
3.4.21.35), submandibular mCK-2 (fragment) [Mus musculus], full
insert sequence.
AK007298
VERSION AK007298.1 GI:12840752
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
1 Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
10349636
2
3 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
11042159
4
5 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsuana, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
11076861
6
7 The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
8
9 The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
10
11 (bases 1 to 450)
12 Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H.,
Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y.,
Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K.,
Hiraoka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M.,
Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M.,
Koya, S., Kurihara, K., Matsuyama, T., Miyazaki, A., Nishi, K.,
Nomura, K., Numasaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C.,
Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D.,
Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y.,
Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T.,
Tejima, Y., Toya, T., Yamamura, T., Yasunishi, A., Yoshida, K.,
Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp,
URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)
Please visit our web site (<http://genome.gsc.riken.jp/>) for further
details.
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues. First strand cDNA was primed with a primer
[5'-GAGAGAGAGAGATCCAGAGCTCTTTTCTTTTCTTNN 3'], cDNA was
prepared by using trehalose thermo-activated reverse transcriptase
and subsequently trehalose for full-length by cap-trapper. Second
strand cDNA was prepared with the primer adapter of sequence [5',
GAGAGAGAGCGCCGCAATTATTCGATTAAATTAATTAATCCCCCCC 3']. cDNA
was cleaved with XhoI and SstI. Cloning sites, 5' end: XhoI; 3'
end: SstI. Host: SOLR.
Location/Qualifiers

[illegible]

Query Match	93.6%	Score 23.4;	DB 5;	Length 460;
Best Local Similarity	96.0%	Pred. No. 12;		
Matches	24: Conservative	0: Mismatches	1: Indels	

QY 1 AAAAAACACTTGGAAAGGAGACTCA 25

Db 230 AAAAACAACCTTGAAGGAGACTCA 254

RESULT 3
BX631074/c
LOCUS
DEFINITION
BX631074 pBluescript Lion Mus musculus cDNA clone LIONp462B0412 3',
mRNA sequence.
BX631074
BX631074.1 GI:33610946
ACCESSION
VERSION
KEYWORDS
SOURCE
Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 298)
REFERENCE
AUTHORS
Henrich, J., Hermann, J., Kranz, H., Loebbert, R., Schluter, T.,
Schuetz, D., Weindel, M., Heil, O., Ebert, L., Neubert, P., Peters, M.,
Radloff, U., Schneider, D., and Korn, B.
TITLE
JOURNAL
COMMENT
Mouse ArrayTAG cDNA (LION)
Contact: Ina Rolfs
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
RZPD; LIONp462B0412.
RZPDLib;
Mouse ArrayTAG cDNA (LION)
http://www.rzpd.de/cgi-bin/products/showLib.pl.cgi/response?libNo=4
62 Contact: Ina Rolfs
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Tel: +49 30 32639 101
Fax: +49 30 32639 111
www.rzpd.de
This clone is available royalty-free from RZPD;
contact RZPD (clone@rzpd.de) for further information. Seq primer:
RP: CAGGAACACTATGAC.

FEATURES
Location/Qualifiers
1..298
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="LIONp462B0412"
/lab_host="DH10B"
/clone_lib="pBluescript Lion"

ORIGIN
Query Match 83.2%; Score 20.8; DB 5; Length 298;
Best Local Similarity 91.7%; Pred. No. 1.7e+02;
Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AAAAACAACCTTGAAGGAGACTCA 25
|||||
Db 218 AAAGACACTTGCNAGGAGACTCA 195
|||||

RESULT 4
BB796159
LOCUS
DEFINITION
BB796159 RIKEN full-length enriched, Leydig cells CRL-2065 MLTC-1
cDNA Mus musculus cDNA clone G4D0006D17 3', mRNA sequence.
BB796159
BB796159.1 GI:16965778
ACCESSION
VERSION
KEYWORDS
SOURCE
Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 464)
REFERENCE
AUTHORS
Akimura, T., Arakawa, T., Carninci, P., Furuno, M., Hanagaki, T.,
Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Imotani, K.,

Ishii, Y., Ito, M., Kawai, J., Kojima, Y., Konno, H., Kouda, M.,
Matsuyama, T., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R.,
Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sakazume, N.,
Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Suzuki, H., Takahashi, A., Takahashi, F., Takaku-Akahira, S.,
Tanaka, T., Tomaru, A., Toya, T., Wataniki, A., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.
RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura, T., et al.
2001)
Unpublished (2001)
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Sushiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gs.c.riken.jp, URL: http://genome.gsc.riken.jp/
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
wagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,
Wataniki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T.,
Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A.
and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)
Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P.,
Sugahara, Y. and Hayashizaki, Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Please visit our web site (http://genome.gsc.riken.go.jp) for
further details.
e mouse tissues.
Location/Qualifiers
1..464
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="G4D0006D17"
/tissue_type="testis"
/cell_type="Leydig cells"
/cell_lines="CRL-2065 MLTC-1"
/clone_lib="RIKEN full-length enriched, Leydig cells
CRL-2065 MLTC-1 cDNA"

ORIGIN
Query Match 83.2%; Score 20.8; DB 2; Length 464;
Best Local Similarity 91.7%; Pred. No. 1.8e+02;
Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AAAAACAACCTTGAAGGAGACTCA 25
|||||
Db 254 AAAGACACTTGTAGGAGACTCA 277
|||||

RESULT 5
BF454696
LOCUS
DEFINITION
BF454696 mao08c01.y1 NCI_CGAP_SG1 Mus musculus cDNA clone IMAGE:3810409 5'-
similar to SW:K1XL_MOUSE P15948 GLANDULAR KALLIKREIN K22 PRECURSOR
;. mRNA sequence.
BF454696
BF454696.1 GI:11520865
ACCESSION
VERSION
KEYWORDS
SOURCE
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Euthera; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Mus.

1 (bases 1 to 466)

NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: cgaps-r@mail.nih.gov

Tissue Procurement: Jeffrey E. Green, M.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

MGI:1452521

Seq primer: -40RP from Gibco

High quality sequence stop: 418.

Location/Qualifiers

1. .466

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

/clone="IMAGE:3810409"

/sex="male"

/tissue_type="salivary gland"

/dev_stage="5 months"

/lab_host="DH10B (T1 phage-resistant)"

/clone_lib="NCI CGAP SGI"

/notes="vector: PCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.6 Kb. Library constructed by Life Technologies."

ORIGIN

Query Match 83.2%; Score 20.8; DB 2; Length 466;

Best Local Similarity 91.7%; Pred. No. 1.8e+02;

Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AAAAACAACCTTGGAGGGAGACTCA 25

DB 314 AAAGACACTTGTAGGGAGACTCA 337

RESULT 6

CL346540

LOCUS

DEFINITION

CL346540 623 bp DNA linear GSS 19-AUG-2004

RPIC144_270F2.r RPCI-44 Sus scrofa genomic clone RPC144_270F2, genomic survey sequence.

ACCESSION

CL346540

VERSION

CL346540.1

KEYWORDS

GSS.

SOURCE

Sus scrofa (pig)

ORGANISM

Sus scrofa

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Euthera; Laurasiatheria; Cetartiodactyla; Suina; Suidae;

Sus.

1 (bases 1 to 623)

Rogatcheva, M.B., Meyers, S., He, W., Larkin, D.M., Marron, B.M.,

Bever, J.B. and Schook, L.B.

Piggy-BACing the Human Genome: Constructing a Porcine Physical Map

Through Comparative Genomics

Unpublished (2004)

Other GSSs: RPC144_270F2.f

Contact: Lawrence B. Schook

Department of Animal Sciences

University of Illinois at Urbana Champaign

1201 W. Gregory Dr., Urbana, IL 61801, USA

Tel: 217 265 5326

Fax: 217 244 5617

Email: schook@uiuc.edu

Clones are derived from the porcine BAC library RPCI-44

(<http://www.bacpac.chori.org/porcine242.htm>). For BAC library availability, please contact Pieter de Jong (pdejong@chori.org). Clones may be purchased from BACPAC Resources (<http://BACPACorders.chori.org>). This work was undertaken as part of the International Swine Genome Sequencing Consortium by the University of Illinois at Urbana Champaign, USA with funds provided by grant No. AG2002-34480-11828 from USDA-CSREES and AG2001-35205-09965 from USDA/NRI (Livestock Genome Sequencing Initiative)

Plate: 270 row: F column: 2

Seq primer: SP6

Class: BAC ends.

Location/Qualifiers

1. .623

/organism="Sus scrofa"

/mol_type="genomic DNA"

/strain="four pigs (breed: 37.5% Yorks Landrace and 25% Meishan)"

/db_xref="taxon:9823"

/clone="RPC144_270F2"

/sex="male"

/cell_type="blood"

/clone_lib="RPCI-44"

/note="vector: pTARBAC2; Site 1: EcoRI; Site 2: EcoRI;

porcine male BAC library produced by Pieter de Jong"

ORIGIN

Query Match 83.2%; Score 20.8; DB 10; Length 623;

Best Local Similarity 91.7%; Pred. No. 1.9e+02;

Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AAAAACAACCTTGGAGGGAGACTCA 25

DB 258 AAAAACAACATGGAGGGAGACTCA 281

RESULT 7

BF454299

LOCUS

DEFINITION

BF454299 666 bp mRNA linear EST 01-DEC-2000

maa03c06.y1 NCI CGAP_SG1 Mus musculus cDNA clone IMAGE:3810034 5', similar to SW:KIK9_MOUSE P15949 GLANDULAR KALLIKREIN K9 PRECURSOR

;; mRNA sequence.

ACCESSION

BF454299

VERSION

BF454299.1

KEYWORDS

EST.

SOURCE

Mus musculus (house mouse)

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Euthera; Euarchontoglires; Glires; Rodentia;

Sciurognathi; Muroidea; Muridae; Murinae; Mus.

1 (bases 1 to 666)

NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

Other ESTs: maa03c06.x1

Contact: Robert Strausberg, Ph.D.

Email: cgaps-r@mail.nih.gov

Tissue Procurement: Jeffrey E. Green, M.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

MGI:1452146

Seq primer: -40RP from Gibco

High quality sequence stop: 439.

Location/Qualifiers

1. .666

/organism="Mus musculus"

/mol_type="mRNA"

```
/db_xref=taxon:10090"
/clone="IMAGE:3810034"
/sex="male"
/tissue_type="salivary gland"
/dev_stage="5 months"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI CGAP SG1"
/note="Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI;
Cloned unidirectionally. Primer: Oligo dT. Average insert
size 1.6 kb. Library constructed by Life Technologies."

ORIGIN
Query Match      83.2%; Score 20.8; DB 2; Length 666;
Best Local Similarity 91.7%; Pred. No. 1.9e+02;
Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AAAAACTTGAAGGAGACTCA 25
Db 643 AAAGACACTTGAAGGAGACTCA 666

RESULT 8
BF538629      669 bp      mRNA      linear      EST 11-DEC-2000
LOCUS        602052947F1 NCI_CGAP_SG2 Mus musculus cDNA clone IMAGE:4192307 5',
DEFINITION   mRNA sequence.
ACCESSION    BF538629
VERSION      BF538629.1 GI:11625997
KEYWORDS
SOURCE       Mus musculus (house mouse)
ORGANISM     Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.
1 (bases 1 to 669)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLM9522 row: m column: 12
High quality sequence stop: 667.
Location/Qualifiers
1. 669
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="IMAGE:4192307"
/clone="IMAGE:4192307"
/clone_lib="NCI CGAP_SG2"
/note="Organ: salivary gland; Vector: pCMV-SPORT6; Site 1:
NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo
dT. Average insert size 1.3 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

ORIGIN
Query Match      83.2%; Score 20.8; DB 2; Length 669;
Best Local Similarity 91.7%; Pred. No. 1.9e+02;
Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AAAAACTTGAAGGAGACTCA 25
Db 450 AAAGACACTTGAAGGAGACTCA 473

RESULT 9
BG872391      680 bp      mRNA      linear      EST 29-MAY-2001
LOCUS        60279684F1 NCI_CGAP_SG2 Mus musculus cDNA clone IMAGE:4923716 5',
DEFINITION   mRNA sequence.
ACCESSION    BG872391
VERSION      BG872391.1 GI:14222931
KEYWORDS
SOURCE       Mus musculus (house mouse)
ORGANISM     Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.
1 (bases 1 to 680)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLM10845 row: h column: 21
High quality sequence stop: 678.
Location/Qualifiers
1. 680
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4923716"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI_CGAP_SG2"
/note="Organ: salivary gland; Vector: pCMV-SPORT6; Site 1:
NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo
dT. Average insert size 1.3 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

ORIGIN
Query Match      83.2%; Score 20.8; DB 2; Length 680;
Best Local Similarity 91.7%; Pred. No. 2e+02;
Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AAAAACTTGAAGGAGACTCA 25
Db 639 AAAGACACTTGAAGGAGACTCA 662

RESULT 10
BG865082      682 bp      mRNA      linear      EST 29-MAY-2001
LOCUS        602784236F1 NCI_CGAP_SG2 Mus musculus cDNA clone IMAGE:4910264 5',
DEFINITION   mRNA sequence.
ACCESSION    BG865082
VERSION      BG865082.1 GI:14215620
KEYWORDS
SOURCE       Mus musculus (house mouse)
ORGANISM     Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.
1 (bases 1 to 682)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
```

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10810 row: h column: 09
High quality sequence stop: 679.
Location/Qualifiers

FEATURES

source

1. .682
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4910264"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI_CGAP_SG2"
/note="Organ: salivary gland; Vector: pCMV-SPORT6; Site 1:
NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo
dT. Average insert size 1.3 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."
83.2%; Score 20.8; DB 2; Length 682;
Best Local Similarity 91.7%; Pred. NO. 2e+02;
Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

ORIGIN

Query Match 83.2%; Score 20.8; DB 2; Length 682;
Best Local Similarity 91.7%; Pred. NO. 2e+02;
Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY

2 AAAAAGACTTGGAGGGAGACTCA 25
|||||

Db

461 AAAAGACTTGGAGGGAGACTCA 484
|||||

RESULT 11

BG869198

LOCUS

DEFINITION 6027899082F1 NCI_CGAP_SG2 Mus musculus cDNA clone IMAGE:4920164 5',
mRNA sequence.

ACCESSION

BG869198

VERSION

BG869198.1

KEYWORDS

EST.

SOURCE

Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.
1 (bases 1 to 683)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10836 row: d column: 21
High quality sequence stop: 660.
Location/Qualifiers

FEATURES

source

1. .683
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4920164"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI_CGAP_SG2"
/note="Organ: salivary gland; Vector: pCMV-SPORT6; Site 1:
NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo
dT. Average insert size 1.3 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

ORIGIN

Query Match 83.2%; Score 20.8; DB 2; Length 683;
Best Local Similarity 91.7%; Pred. NO. 2e+02;
Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY

2 AAAAAGACTTGGAGGGAGACTCA 25
|||||

Db

452 AAAAGACTTGTAGGGAGACTCA 475
|||||

RESULT 12

BG865647

LOCUS

DEFINITION 602783810F1 NCI_CGAP_SG2 Mus musculus cDNA clone IMAGE:4909778 5',
mRNA sequence.

ACCESSION

BG865647

VERSION

BG865647.1

KEYWORDS

EST.

SOURCE

Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.
1 (bases 1 to 737)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10809 row: d column: 03
High quality sequence stop: 736.
Location/Qualifiers

FEATURES

source

1. .737
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4909778"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI_CGAP_SG2"
/note="Organ: salivary gland; Vector: pCMV-SPORT6; Site 1:
NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo
dT. Average insert size 1.3 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

ORIGIN

Query Match 83.2%; Score 20.8; DB 2; Length 737;
Best Local Similarity 91.7%; Pred. NO. 2e+02;
Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY

2 AAAAAGACTTGGAGGGAGACTCA 25
|||||

Db

643 AAAAGACTTGTAGGGAGACTCA 666
|||||

RESULT 13

BG866078

LOCUS

DEFINITION 602787987F1 NCI_CGAP_SG2 Mus musculus cDNA clone IMAGE:4914236 5',
mRNA sequence.

ACCESSION

BG866078

VERSION

BG866078.1

KEYWORDS

EST.

SOURCE

Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.

Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AAAAACAACCTTGAAGGAGACTC 24
Db 105 AAAAAGCACTTGAAGGAGACTC 128

RESULT 16
BG870364
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

775 bp mRNA linear EST 29-MAY-2001
602791257F1 NCI_CGAP_SG2 Mus musculus cDNA clone IMAGE:4922391 5',
mRNA sequence.
BG870364
EST.
Mus musculus (house mouse)

1. .856
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4914211"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI CGAP SG2"
/note="Organ: salivary gland; Vector: pCMV-SPORT6; Site: 1:
NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo
dT. Average insert size 1.3 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-rc@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LHAM10820 row: 1 column: 20
High quality sequence stop: 759.
Location/Qualifiers

1. .856
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4914211"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI CGAP SG2"
/note="Organ: salivary gland; Vector: pCMV-SPORT6; Site: 1:
NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo
dT. Average insert size 1.3 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

ORIGIN
Query Match 83.2%; Score 20.8; DB 2; Length 856;
Best Local Similarity 91.7%; Pred. No. 2e+02; Indels 0; Gaps 0;
Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 AAAAACAACCTTGAAGGAGACTCA 25
Db 630 AAAGACACTTGAAGGAGACTCA 653

RESULT 18
BF302440
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

881 bp mRNA linear EST 21-NOV-2000
602031451F1 NCI_CGAP_SG2 Mus musculus cDNA clone IMAGE:4166413 5',
mRNA sequence.
BF302440
EST.
Mus musculus (house mouse)

1. .881
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4166413"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI CGAP SG2"
/note="Organ: salivary gland; Vector: pCMV-SPORT6; Site: 1:
NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo
dT. Average insert size 1.3 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-rc@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LHAM9455 row: f column: 14
High quality sequence stop: 761.
Location/Qualifiers

1. .881
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4166413"

FEATURES
source

AUTHORS
TITLE
JOURNAL
COMMENT

NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-rc@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LHAM10820 row: 1 column: 20
High quality sequence stop: 759.
Location/Qualifiers

1. .856
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4914211"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI CGAP SG2"
/note="Organ: salivary gland; Vector: pCMV-SPORT6; Site: 1:
NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo
dT. Average insert size 1.3 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

ORIGIN
Query Match 83.2%; Score 20.8; DB 2; Length 856;
Best Local Similarity 91.7%; Pred. No. 2e+02; Indels 0; Gaps 0;
Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 AAAAACAACCTTGAAGGAGACTCA 25
Db 630 AAAGACACTTGAAGGAGACTCA 653

RESULT 18
BF302440
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

881 bp mRNA linear EST 21-NOV-2000
602031451F1 NCI_CGAP_SG2 Mus musculus cDNA clone IMAGE:4166413 5',
mRNA sequence.
BF302440
EST.
Mus musculus (house mouse)

1. .881
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4166413"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI CGAP SG2"
/note="Organ: salivary gland; Vector: pCMV-SPORT6; Site: 1:
NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo
dT. Average insert size 1.3 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-rc@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LHAM9455 row: f column: 14
High quality sequence stop: 761.
Location/Qualifiers

1. .881
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4166413"

FEATURES
source

http://image.llnl.gov
 Plate: LLAM9459 row: f column: 18
 High quality sequence stop: 830.
 Location/Qualifiers

FEATURES

source
 1..952
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:4167953"
 /lab_host="DH10B (T1 phage-resistant)"
 /clone_lib="NCI CGAP SG2"
 /notes="Organ: salivary gland; Vector: pCMV-SPORT6; Site 1:
 Note1: Site 2: SalI; Cloned unidirectionally. Primer: Oligo
 dt. Average insert size 1.3 kb. Constructed by Life
 Technologies. Note: this is a NCI_CGAP Library."

ORIGIN

Query Match 83.2%; Score 20.8; DB 2; Length 952;
 Best Local Similarity 91.7%; Pred. No. 2.1e+02;
 Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AAAAAACACTTGAAGGGAGACTCA 25
 |||||
 Db 616 AAAGACACTTGTAAAGGGAGACTCA 639

RESULT 22
 BZ023877/c
 LOCUS
 DEFINITION
 711 bp DNA linear GSS 08-OCT-2002
 oej6901.b1 B.oleracea002 Brassica oleracea genomic, genomic survey
 sequence.

ACCESSION BZ023877
 VERSION BZ023877.1 GI:23593611
 KEYWORDS
 SOURCE

ORGANISM

Brassica oleracea
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
 1 (bases 1 to 711)
 Delehaunty, K., Fewell, G., Fulton, L., McCombie, W.R., Miner, T.,
 Nash, W., Rabinowicz, P.D. and Wilson, R.K.
 Whole genome shotgun reads from Brassica oleracea
 Unpublished (2002)
 Contact: Richard K. Wilson
 Genome Sequencing Center
 Washington University School of Medicine
 Email: submissions@wustl.wustl.edu
 Plate: oej69 row: g column: 01
 Seq primer: -21uppt forward
 Class: shotgun
 High quality sequence start: 16
 High quality sequence stop: 529.

FEATURES

source
 1..711
 /organism="Brassica oleracea"
 /mol_type="genomic DNA"
 /db_xref="taxon:3712"
 /clone_lib="B.oleracea002"
 /note="Vector: pOTw13; Whole genome shotgun library from
 flowering buds. DNA was purified from a crude nuclear
 prep using Brassica oleracea T01000DH3 buds provided by
 Thomas Osborn at the University of Wisconsin. Genomic
 DNA was provided by Pablo Rabinowicz (CSHL) and the
 shotgun library prepared at Washington University Genome
 Sequencing Center."

ORIGIN

Query Match 80.8%; Score 20.2; DB 9; Length 711;
 Best Local Similarity 88.0%; Pred. No. 3.7e+02;
 Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAACACTTGAAGGGAGACTCA 25
 |||||
 Db 122 AAAAAAATTTGAAGGGAGACTTA 98

RESULT 23

BG866696
 LOCUS

DEFINITION

730 bp mRNA linear

EST. EST 29-MAY-2001

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Mus musculus

(house mouse)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

Sciurognathi; Muridae; Murinae; Mus.

NIH-MGC http://mgc.nci.nih.gov/

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgabbs@mail.nih.gov

Tissue Procurement: Jeffrey E. Green, M.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM10838 row: 1 column: 07

High quality sequence stop: 730.

Location/Qualifiers

1..730

/organism="Mus musculus"

/mol_type="mRNA"

/strain="FVB/N"

/db_xref="taxon:10090"

/clone="IMAGE:4921110"

/lab_host="DH10B (T1 phage-resistant)"

/clone_lib="NCI CGAP SG2"

/note="Organ: salivary gland; Vector: pCMV-SPORT6; Site 1:

Note1: Site 2: SalI; Cloned unidirectionally. Primer: Oligo

dt. Average insert size 1.3 kb. Constructed by Life

Technologies. Note: this is a NCI_CGAP Library."

ORIGIN

Query Match 80.8%; Score 20.2; DB 2; Length 730;

Best Local Similarity 88.0%; Pred. No. 3.7e+02;

Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAACACTTGAAGGGAGACTCA 25

|||||

Db 634 AACAGACACTTGTAAAGGGAGACTCA 658

RESULT 24

BG866696

LOCUS

DEFINITION

719 bp mRNA linear

EST. EST 29-MAY-2001

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Mus musculus

(house mouse)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

Sciurognathi; Muridae; Murinae; Mus.

NIH-MGC http://mgc.nci.nih.gov/

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgabbs@mail.nih.gov

Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Jeffrey E. Green, M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Plate: LLAM10814 row: o column: 22
 High quality sequence stop: 717.
 Location/Qualifiers
 1. .719
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:4911981"
 /lab_host="DH10B (T1 phage-resistant)"
 /clone_lib="NCI CGAP_SG2"
 /note="Organ: salivary gland; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.3 kb. Constructed by Life Technologies. Note: this is a NCI_CGAP Library."

ORIGIN

Query Match 79.2%; Score 19.8; DB 2; Length 719;
 Best Local Similarity 91.3%; Pred. No. 5.6e+02;
 Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 AAAACACTTGGAGGGAGACTCA 25
 |||||||
 Db 638 AAGACACTTGTAAAGGGAGACTCA 660

FEATURES
 source

RESULT 25
 BF534137
 LOCUS
 DEFINITION
 602047621F1 NCI_CGAP_SG2 Mus musculus cDNA clone IMAGE:4187212 5', mRNA sequence.
 BF534137
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Mus musculus (house mouse)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Mus.
 1 (bases 1 to 750)
 NIH-MGC <http://mgc.nci.nih.gov/>
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Jeffrey E. Green, M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Plate: LLAM9509 row: i column: 05
 High quality sequence stop: 718.
 Location/Qualifiers
 1. .750
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:4187212"
 /lab_host="DH10B (T1 phage-resistant)"
 /clone_lib="NCI_CGAP_SG2"

ORIGIN

Query Match 79.2%; Score 19.8; DB 2; Length 794;
 Best Local Similarity 91.3%; Pred. No. 5.7e+02;
 Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 AAAACACTTGGAGGGAGACTCA 25
 |||||||
 Db 629 AAGACACTTGTAAAGGGAGACTCA 651

FEATURES
 source

RESULT 27
 DN792152/c
 LOCUS
 DEFINITION
 DN792152 Sea Urchin primary mesenchyme cell cDNA library Strongylocentrotus purpuratus cDNA clone PMCSPR2-131K22 5', mRNA sequence.


```

ACCESSION DN792152
VERSION DN792152.1 GI:62382219
KEYWORDS EST.
SOURCE Strongylocentrotus purpuratus
ORGANISM Strongylocentrotus purpuratus
Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
Echinoidea; Euechinoidea; Echinacea; Echinoida;
Strongylocentrotidae; Strongylocentrotus.

REFERENCE 1 (bases 1 to 878)
AUTHORS Zhu,X., Mahairas,G., Illies,M.R., Cameron,R.A., Davidson,E.H. and
Ettensohn,C.A.
TITLE A large-scale analysis of mRNAs expressed by primary mesenchyme
cells of the sea urchin embryo
JOURNAL Development 128 (13), 2615-2627 (2001)
PUBMED 11493577
COMMENT Contact: Erica Sodergren
Human Genome Sequencing Center
Baylor College of Medicine
One Baylor Plaza, Houston, TX 77030, USA
Tel: 713-798-7676
Fax: 713-798-6977
Email: ericas@bcm.tmc.edu
NCBI Trace Archive: 490875627
Insert Length: 1750 Std Error: 0.25
Plate: 131 row: K column: 22.

FEATURES             source
    Location/Qualifiers
        1..878
        /organism="Strongylocentrotus purpuratus"
        /mol_type="mRNA"
        /db_xref="taxon:7668"
        /clone="PMCSRP2-131K22"
        /tissue_type="embryo"
        /cell_type="primary mesenchyme cells"
        /lab_host="E.coli"
        /clone_lib="Sea Urchin primary mesenchyme cell cDNA
        library"
        /note="Vector: pSPORT1; Site 1: NotI; Site 2: MluI; oligo
        dt priming from poly A+ RNA, directionally cloned"

ORIGIN
Query Match          79.2%; Score 19.8; DB 8; Length 878;
Best Local Similarity 91.3%; Pred. No. 5.8e+02;
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AAAAAACACTGGAGGAGACT 23
Db 749 AAAAAACACATGGAGGAGACT 727

RESULT 28
BF300019 925 bp mRNA linear EST 21-NOV-2000
LOCUS 602030681F1 NCI_CGAP_SG2 Mus musculus cDNA clone IMAGE:4165557 5',
DEFINITION mRNA sequence.
ACCESSION BF300019
VERSION BF300019.1 GI:11246542
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 925)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: gcgaps-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM9453 row: b column: 22
High quality sequence stop: 698.

FEATURES             source
    Location/Qualifiers
        1..925
        /organism="Mus musculus"
        /mol_type="mRNA"
        /strain="FVB/N"
        /db_xref="taxon:10090"
        /clone="IMAGE:4165557"
        /lab_host="DH10B (T1 phage-resistant)"
        /clone_lib="NCI_CGAP_SG2"
        /note="Organ: salivary gland; Vector: pCMV-SPORT6; Site 1:
        NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo
        dt. Average insert size 1.3 kb. Constructed by Life
        Technologies. Note: this is a NCI_CGAP Library."

ORIGIN
Query Match          79.2%; Score 19.8; DB 2; Length 925;
Best Local Similarity 91.3%; Pred. No. 5.8e+02;
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 AAAAACTTGGAGGAGACTC 24
Db 610 AAAGACACTGTAAGGGAGACTC 632

RESULT 29
DN561586/c
LOCUS DN561586/c
DEFINITION Strongylocentrotus purpuratus cDNA clone PMCSRP2-101M22 5', mRNA
sequence.
ACCESSION DN561586
VERSION DN561586.1 GI:61120625
KEYWORDS EST.
SOURCE Strongylocentrotus purpuratus
ORGANISM Strongylocentrotus purpuratus
Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
Echinoidea; Euechinoidea; Echinacea; Echinoida;
Strongylocentrotidae; Strongylocentrotus.

REFERENCE 1 (bases 1 to 963)
AUTHORS Zhu,X., Mahairas,G., Illies,M.R., Cameron,R.A., Davidson,E.H. and
Ettensohn,C.A.
TITLE A large scale analysis of mRNAs expressed by primary mesenchyme
cells of the sea urchin embryo
JOURNAL Development 128 (13), 2615-2627 (2001)
PUBMED 11493577
COMMENT Contact: Erica Sodergren
Human Genome Sequencing Center
Baylor College of Medicine
One Baylor Plaza, Houston, TX 77030, USA
Tel: 713-798-7676
Fax: 713-798-6977
Email: ericas@bcm.tmc.edu
NCBI Trace Archive: 486717752
Insert Length: 1750 Std Error: 0.25
Plate: 101 row: M column: 22.

FEATURES             source
    Location/Qualifiers
        1..963
        /organism="Strongylocentrotus purpuratus"
        /mol_type="mRNA"
        /db_xref="taxon:7668"
        /clone="PMCSRP2-101M22"
        /tissue_type="embryo"
        /cell_type="primary mesenchyme cells"
        /lab_host="E.coli"
        /clone_lib="Sea Urchin primary mesenchyme cell cDNA
        library"
        /note="Vector: pSPORT1; Site 1: NotI; Site 2: MluI; oligo
        dt priming from poly A+ RNA, directionally cloned"

ORIGIN

```

Query Match 79.2%; Score 19.8; DB 8; Length 963;
Best Local Similarity 91.3%; Pred. No. 5.8e+02;
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAACACTTGAAGGAGACT 23
|||||
Db 209 AAAAAACATGGAGGAGACT 187

RESULT 30
DN661091
LOCUS
DEFINITION
CEC50-E06.yld-s SHGC-CEC Gasterosteus aculeatus cDNA clone
CEC50-E06 5', mRNA sequence.
DN661091
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Gasterosteus aculeatus (three spined stickleback)
Gasterosteus aculeatus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Gasterosteiformes;
Gasterosteidae; Gasterosteus.
1 (bases 1 to 1203)
Kingsley,D.M., Peichel,C., Balabhadra,S., Grimwood,J., Dickson,M.,
Schmutz,J. and Myers,R.M.
Expressed sequence tags from Gasterosteus aculeatus
Unpublished (2003)
Contact: Grimwood, Jane
Stanford Human Genome Center
Stanford University School of Medicine
975 S California Ave, Palo Alto, CA 94304, USA
Tel: 650 320 5917
Fax: 650 320 5801
Email: jane@shgc.stanford.edu
Plate: 50
High quality sequence stop: 711.

Location/Qualifiers
1. .1203
/organism="Gasterosteus aculeatus"
/mol_type="mRNA"
/strain="Conner Creek sticklebacks, WA USA"
/db_xref="taxon:69293"
/clone="CEC50-E06"
/sex="mixed male and female"
/tissue_type="skin"
/dev_stage="adult"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="SHGC-CEC"
/notes="Vector: Express 1; Total and poly A+ RNA was isolated from the indicated stickleback tissue, and a cDNA library was constructed in the Express 1 plasmid vector by Open Biosystems. First strand cDNA synthesis was primed with an 54 bp linker primer containing an oligodT sequence preceded by a synthetic NotI site (first strand primer: 5'-GACTAGTCTAGATCGGCGCGCC(T)25-3'). Following second strand synthesis, cDNAs were made blunt at the end corresponding to the original 5 prime end of mRNA, and cloned directionally into the NotI and EcoRV sites of Express 1. Note that the EcoRV site is typically destroyed in the blunt end cloning, leaving a junction of the form 'xxxATC' (where ATC is the second half of the EcoRV site, and xxx is derived from the cDNA sequence). A map of the Express 1 vector is available at:
<http://www.openbiosystems.com/cdna.library.constructionfaq.php#8>
The primary library was transformed and amplified in DH10B (T1 phage resistant) bacteria. Clones available from Open Biosystems:
<http://www.openbiosystems.com/stickleback>"

FEATURES
source

Query Match 79.2%; Score 19.8; DB 8; Length 1203;
Best Local Similarity 91.3%; Pred. No. 5.8e+02;
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Best Local Similarity 91.3%; Pred. No. 6e+02;
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AAAAAACACTTGAAGGAGACTC 24
|||||
Db 1125 AAAAAAACTTGAAGGAAACTC 1147

RESULT 31
CW916421/c
LOCUS
DEFINITION
RPC142_160K4.TJ RPC1-42 Bos taurus genomic clone RPC142_160K4,
genomic survey sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Bos taurus (cow)
Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
Pecora; Bovidae; Bovinae; Bos.

REFERENCE
AUTHORS
1 (bases 1 to 411)
Larkin,D.M., Donthu,K., LeDuc,R., Ryan,K., Liu,L., de Jong,P.J. and
Lewin,H.A.

TITLE
JOURNAL
COMMENT
End sequencing of Holstein BAC library RPC1-42
Unpublished (2004)
Other_GSSs: RPC142_160K4.TV
Contact: Harris Lewin

Department of Animal Sciences
University of Illinois at Urbana Champaign
1201 W. Gregory Dr., Urbana, IL 61801, USA
Tel: 217 333 5998
Fax: 217 244 5617
Email: h-lewin@uiuc.edu
Clones are derived from the bovine BAC library RPC1-42
(<http://bacpac.choi.org/mbovine42.htm>). For BAC library availability, please contact Pieter de Jong
(pdejong@mail.choi.org). Clones may be purchased from BACPAC
Resources (http://www.choi.org/bacpac/ordering_information.htm).
Funding was provided in part by grant no. AG2004-34480-14417 from
USDA-CSREES (Livestock Genome Sequencing Initiative) and
AG58-5438-2-313 from USDA-ARS
Plate: 160 row: K column: 4
Seq primer: SP6
Class: BAC ends.

Location/Qualifiers
1. .411
/organism="Bos taurus"
/mol_type="genomic DNA"
/strain="Holstein"
/db_xref="taxon:9913"
/clone="RPC142_160K4"
/sex="male"
/cell_type="blood"
/clone_lib="RPC1-42"
/note="Vector: pBACe3.6; Site 1: EcoRI; Site 2: EcoRI;
RPC1-42 Bovine BAC library (Male) produced by Pieter de
Jong"

FEATURES
source

Query Match 77.6%; Score 19.4; DB 10; Length 411;
Best Local Similarity 95.2%; Pred. No. 7.8e+02;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

ORIGIN

QY 5 AACACTTGAAGGAGACTCA 25
|||||
Db 282 AACACTTAGAGGAGACTCA 262

RESULT 32

CNS02ES8
LOCUS
DEFINITION

CNS02ES8 745 bp DNA linear GSS 01-SEP-2000
Tetraodon nigroviridis genome survey sequence T7 end of clone

131E09 of library G from Tetraodon nigroviridis, genomic survey sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AL194129
AL194129.1 GI:7832235
GSS; genome survey sequence.
Tetraodon nigroviridis

Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percormorpha; Tetraodontiformes;
Tetraodontidae; Tetraodontidae; Tetraodon.

REFERENCE
AUTHORS

1
Roest Crolius.H., Jaillon.O., Dasilva.C., Bouneau.L., Fisher.C.,
Bernot.A., Fizes.C., Wincker.P., Brottier.P., Quetier.F.,
Saurin.W. and Weissenbach.J.

Estimate of human gene number provided by genome-wide analysis
using Tetraodon nigroviridis DNA sequence

JOURNAL
PUBMED

Nat. Genet. 25 (2), 235-238 (2000)

REFERENCE
AUTHORS

2
Roest Crolius.H., Jaillon.O., Dasilva.C., Ozouf-Costaz.C.,
Fizes.C., Fischer.C., Bouneau.L., Billault.A., Quetier.F.,
Saurin.W., Bernot.A. and Weissenbach.J.

Characterization and repeat analysis of the compact genome of the
freshwater pufferfish Tetraodon nigroviridis

JOURNAL
PUBMED

Genome Res. 10 (7), 939-949 (2000)

REFERENCE
AUTHORS

3 (bases 1 to 745)

Direct Submission
Submitted (12-APR-2000) Genoscope - Centre National de Sequencage :

BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr
- Web : www.genoscope.cns.fr)

This sequence is a single read and was generated as part of a large
scale clone-end sequencing project of the Tetraodon nigroviridis
genome. For more information, please take a look at
http://www.genoscope.cns.fr/tetraodon.

FEATURES
source

1..745
Location/Qualifiers
/organism="Tetraodon nigroviridis"
/mol_type="genomic DNA"
/db_xref="taxon:99883"
/clone="131E09"
/clone_lib="G"
/note="Genoscope sequence ID : COAG131AC05LP1
end : 77"

ORIGIN

Query Match 77.6%; Score 19.4; DB 10; Length 745;
Best Local Similarity 95.2%; Pred. No. 8.6e+02;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AAAAACAACCTTGGGAAGGAGA 21

Db 238 AAAAACAACCTTGGGAAGGAGA 258

RESULT 33
LOCUS

BI283322/c
DEFINITION
UI-R-DA0-bzf-g-06-0-UI.s1 UI-R-DA0 Rattus norvegicus cDNA clone
UI-R-DA0-bzf-g-06-0-UI 3', mRNA sequence.

ACCESSION
VERSION

BI283322
BI283322.1 GI:14934954

KEYWORDS
SOURCE

Rattus norvegicus (Norway rat)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE
AUTHORS

1 (bases 1 to 303)
Ronald.M.F., Lennon.G. and Soares.M.B.

Normalization and subtraction: two approaches to facilitate gene

JOURNAL
PUBMED
COMMENT

discovery
Genome Res. 6 (9), 791-806 (1996)

8889548

Contact: Soares, MB
Coordinated Laboratory for Computational Genomics

University of Iowa

375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA

Tel: 319 335 8250

Fax: 319 335 9565

Email: bento-soares@uiowa.edu

The sequence contained an oligo-dT track that was present in the
oligonucleotide that was used to prime the synthesis of first
strand cDNA and therefore this may represent a bonafide poly A
tail. The sequence tag present in the cDNA between the NotI site
and the oligo-dT track served to verify it as a clone from the
non-normalized rat eye library cDNA Library Preparation: M.B.
Soares Lab Clone distribution: clones will be available through
Research Genetics (www.resgen.com)

Seq primer: M13 Forward

POLYA=Yes.

FEATURES
source

Location/Qualifiers

1..303

/organism="Rattus norvegicus"

/mol_type="mRNA"

/strain="Sprague-Dawley"

/db_xref="taxon:10116"

/clone="UI-R-DA0-bzf-g-06-0-UI"

/dev_stage="ADULT"

/lab_host="DH10B (Life Technologies)"

/clone_lib="UI-R-DA0"

/notes="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; The UI-R-DA0
library is a non-normalized library constructed from rat
salivary gland tissue. For a detailed description of the
library from which this clone was derived, please visit
our web site at ratest.eng.uiowa.edu. The subtraction has
been previously described in (Bonaldo, Lennon and Soares,
Genome Research 6:791-806, 1996)

TAG TISSUE=rat eye

TAG_LIB=UI-R-DA0

TAG_SEQ=CAGCC"

ORIGIN

Query Match 76.8%; Score 19.2; DB 2; Length 303;
Best Local Similarity 87.5%; Pred. No. 9.2e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 AAAAACAACCTTGGGAAGGAGACTCA 25

Db 234 AAAGACACTTGCACGAGGTGACTCA 211

RESULT 34

BU096127

LOCUS

DEFINITION

tca-612 tca Trypanosoma carassii cDNA clone 03m9 5', mRNA sequence.

ACCESSION

BU096127

VERSION

BU096127.1 GI:25123851

KEYWORDS

EST.

SOURCE

ORGANISM

Trypanosoma carassii

Trypanosoma carassii

Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;

Trypanosoma.

1 (bases 1 to 306)

Agarath.P., Campo.V., Cremona.L., Jager.A., Di Noia,J.M.,

Overath.P., Sanchez,D.O. and Frasch,A.C.

Gene discovery in the freshwater fish parasite Trypanosoma

carassii: identification of trans-sialidase-like and mucin-like

genes

Infect. Immun. 70 (12), 7140-7144 (2002)

Contact: Sanchez DO

Genomics and Bioinformatics

Instituto de Investigaciones Bioteconologicas

Av. Gral Paz S/N, INTI, Edificio 24, B 1650 KNA, San Martin, Buenos Aires, Argentina
Tel: (54-11) 4580/7255/7
Fax: (54-11) 4752-9639
Email: dsanchez@ib.unsam.edu.ar
Sequences were basecalled with phred and vector was masked with crossmatch (see <http://www.phrap.org>). Sequences were then trimmed from both ends to remove low quality bases and masked vector.
Plate: 03 row: m column: 9
Seq primer: R7.

FEATURES
source

Location/Qualifiers
1..306
/organism="Trypanosoma carassii"
/mol_type="mrna"
/db_xref="taxon:38249"
/clone="03m9"
/dev_stage="blood trypanostigote"
/lab_host="Goldfish (Carassius auratus)"
/clone_lib="tca"
/note="Vector: pSport1; Blood trypanostigotes were obtained from goldfish and cultured as described (Overath et al. Parasitol Res (1998) 84:343) before obtaining total RNA using Trizol. cDNA library construction was made from polyA+ mRNA using a poly-dT oligonucleotide as primer. The cDNAs were cloned in a oriented manner using a commercial kit (SuperScript Plasmid System for cDNA Synthesis and Plasmid Cloning, Life Technologies)."

ORIGIN

Query Match 76.8%; Score 19.2; DB 5; Length 306;
Best Local Similarity 87.5%; Pred. No. 9.2e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAACACTTGAAGGAGACTC 24
||||| ||||| ||||| ||||| |||||
Db 54 AAAAAAACTTGAAGGAGAGCG 77

RESULT 35
CR477147/c

LOCUS
DEFINITION
CR477147 Rat pBluescript Lion Rattus norvegicus cDNA clone
LI0NP463H10365 3', mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

CR477147 315 bp mRNA linear EST 07-JUL-2004
Rattus norvegicus (Norway rat)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
1 (bases 1 to 315)

REFERENCE
AUTHORS
Schuette, D., Weindel, M., Heil, O., Ebert, L., Neubert, P., Peters, M., Radelof, U., Schneider, D. and Korn, B.

Rat ArrayTAG cDNA
Unpublished (2004)
Contact: Inge Arlart
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Email: www.rzpd.de
RZPD; LI0NP463H10365.
RZPDLIB;

Rat ArrayTAG cDNA
<http://www.rzpd.de/cgi-bin/products/showlib.pl.cgi?response2libNo=463> Contact: Inge Arlart
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Tel: +49 30 32639 100
Fax: +49 30 32639 111
www.rzpd.de

This clone is available royalty-free from RZPD;
contact RZPD (clone@rzpd.de) for further information. Seq primer:

RP: CAGGAACACGCTATGAC.

FEATURES
source

Location/Qualifiers
1..315
/organism="Rattus norvegicus"
/mol_type="mrna"
/db_xref="taxon:10116"
/clone="LI0NP463H10365"
/lab_host="DH10B"
/clone_lib="Rat pBluescript Lion"

ORIGIN

Query Match 76.8%; Score 19.2; DB 7; Length 315;
Best Local Similarity 87.5%; Pred. No. 9.3e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAACACTTGAAGGAGACTCA 25
||||| ||||| ||||| ||||| |||||
Db 217 AAAGACACTTGCAGGGTGACTCA 194

RESULT 36
CO150776

LOCUS
DEFINITION
CO150776 Aspergillus flavus Normalized cDNA Expression Library
EST825829 Aspergillus flavus cDNA clone NAGEK12 5', end similar to (Q7S1X9)
Aspergillus flavus protein, mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

CO150776 323 bp mRNA linear EST 17-JUN-2004
Aspergillus flavus
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
1 (bases 1 to 323)
Yu, J., Whitelaw, C.A., Nierman, W.C., Bhatnagar, D. and Cleveland, T.E.
Aspergillus flavus expressed sequence tags for identification of
genes with putative roles in aflatoxin contamination of crops
FEMS Microbiol. Lett. (2004) In press
Contact: Yu J
Food and Feed Safety Research Unit
USDA/ARS, Southern Regional Research Center
1100 Robert E. Lee Boulevard, New Orleans, LA 70124, USA
Tel: 504 286 4405
Fax: 504 286 4419
Email: jiuyu@rrrc.ars.usda.gov
Contact Dr. Yu at USDA/ARS SRRC (jiuyu@rrrc.ars.usda.gov) for clone
information
PCR Primers
FORWARD: M13F
BACKWARD: M13R
Seq primer: M13 Forward
POLYA=No.

FEATURES
source

Location/Qualifiers
1..323
/organism="Aspergillus flavus"
/mol_type="mrna"
/strain="NRRL 3357"
/db_xref="taxon:5059"
/clone="NAGEK12"
/sex="asexual mycelia"
/cell_type="mycelia"
/dev_stage="developmental stages from 18 to 96 hours"
/lab_host="E. coli DH10B T1 resistant cells"
/clone_lib="Aspergillus flavus Normalized cDNA Expression Library"
/note="Vector: pBlueScript (SK+) (Stratagene), antibiotic selection marker: Carbenicillin; Site 1: NotI, at the 5 prime end; Site 2: EcoRI, at the 3 prime end; This normalized cDNA expression library was constructed using a mixture of mycelial cells grown under eight different medium conditions and harvested at 5 time points (18, 24, 48, 72, 96 hours). The poly-A sequence was trimmed off before ligating to vector."

ORIGIN

Query Match 76.8%; Score 19.2; DB 7; Length 323;
 Best Local Similarity 87.5%; Pred. No. 9.3e+02;
 Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAACTTGGGAAGGAGACTCA 25
 |||||
 Db 115 AAAAACTTGGGAAGGAGACTCA 138

RESULT 37

B1284068/c 354 bp mRNA linear EST 19-JUL-2001
 LOCUS UI-R-DA0-bzd-h-04-0-UI.81 UI-R-DA0 Rattus norvegicus cDNA clone
 DEFINITION UI-R-DA0-bzd-h-04-0-UI 3', mRNA sequence.

ACCESSION

VERSION B1284068

KEYWORDS

SOURCE B1284068.1 GI:14936396

ORGANISM

Rattus norvegicus (Norway rat)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE

1 (bases 1 to 354)

AUTHORS

Bonaldo, M.F., Lennon, G. and Soares, M.B.

TITLE

Normalization and subtraction: two approaches to facilitate gene

JOURNAL

Genome Res. 6 (9), 791-806 (1996)

PUBMED

8889548

COMMENT

Contact: Soares, MB
 Coordinated Laboratory for Computational Genomics
 University of Iowa
 375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: bento-soares@uiowa.edu
 The sequence contained an oligo-dT track that was present in the
 oligonucleotide that was used to prime the synthesis of first
 strand cDNA and therefore this may represent a bonafide poly A
 tail. The sequence tag present in the cDNA between the NotI site
 and the oligo-dT track served to verify it as a clone from the
 non-normalized rat eye library cDNA Library Preparation: M.B.
 Soares Lab Clone distribution: clones will be available through
 Research Genetics (www.resgen.com)
 Seq primer: M13 Forward
 POLYA=Yes.

FEATURES

source Location/Qualifiers

1..354
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 /clone="UI-R-DA0-bzd-h-04-0-UI"
 /dev_stage="ADULT"
 /lab_host="DH10B (Life Technologies)"
 /clone_lib="UI-R-DA0"
 /note="Vector: p7713D-Pac (Pharmacia) with a modified
 polylinker; Site 1: Not I; Site 2: Eco RI; The UI-R-DA0
 library is a non-normalized library constructed from rat
 salivary gland tissue. For a detailed description of the
 library from which this clone was derived, please visit
 our web site at ratest.eng.uiowa.edu. The subtraction has
 been previously described in (Bonaldo, Lennon and Soares,
 Genome Research 6:791-806, 1996)
 TAG TISSUE=rat eye
 TAG_LIB=UI-R-DA0
 TAG_SEQ=CAGCC"

ORIGIN

Query Match 76.8%; Score 19.2; DB 2; Length 354;
 Best Local Similarity 87.5%; Pred. No. 9.4e+02;
 Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY

2 AAAAACTTGGGAAGGAGACTCA 25
 |||||
 Db 232 AAGACACTTGCAGGGTGACTCA 209

RESULT 38

BX633808/c 368 bp mRNA linear EST 12-AUG-2003
 LOCUS BX633808 pBluescript Lion Mus musculus cDNA clone LIONp462D0812 3',
 DEFINITION mRNA sequence.

ACCESSION

VERSION BX633808.1 GI:33613683

KEYWORDS

SOURCE BX633808.1

ORGANISM

Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 (bases 1 to 368)

AUTHORS

Henrich, J., Hermanns, J., Kranz, H., Loebbert, R., Schlueter, T.,
 Schuette, D., Weindel, M., Heil, O., Ebert, L., Neubert, P., Peters, M.,
 Radloff, U., Schneider, D. and Korn, B.

TITLE

Mouse ArrayTAG cDNA (LION)

JOURNAL

Unpublished (2003)

COMMENT

Contact: Ina Rolf
 RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
 Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
 RZPD; LIONp462D0812.
 RZPDLIB;
 Mouse ArrayTAG cDNA (LION)
<http://www.rzpd.de/cgi-bin/products/showLib.pl.cgi?response?libNo=4>
 62 Contact: Ina Rolf
 RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
 Heubnerweg 6, D-14059 Berlin, Germany
 Tel: +49 30 32639 101
 Fax: +49 30 32639 111
www.rzpd.de
 This clone is available royalty-free from RZPD;
 contact RZPD (clone@rzpd.de) for further information. Seq primer:
 RP: CAGGAACAGCTATGAC.

FEATURES

source Location/Qualifiers
 1..368
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="LIONp462D0812"
 /lab_host="DH10B"
 /clone_lib="pBluescript Lion"

ORIGIN

Query Match 76.8%; Score 19.2; DB 5; Length 368;
 Best Local Similarity 87.5%; Pred. No. 9.5e+02;
 Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY

2 AAAAACTTGGGAAGGAGACTCA 25
 |||||
 Db 217 AAGACACTTGCAGGGTGACTCA 194

RESULT 39

AK006516 375 bp mRNA linear HTC 03-APR-2004
 LOCUS AK006516 adult male testis cDNA, RIKEN full-length enriched
 DEFINITION library, clone:1700029022 product:nerve growth factor, gamma, full
 insert sequence.

ACCESSION

VERSION AK006516.1 GI:12839658

KEYWORDS

SOURCE AK006516

ORGANISM

Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

REFERENCE	Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
AUTHORS	1 (bases 1 to 391) Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE	Normalization and subtraction: two approaches to facilitate gene discovery
JOURNAL	Genome Res. 6 (9), 791-806 (1996)
PUBMED	8889548
COMMENT	Contact: Soares, MB Coordinated Laboratory for Computational Genomics University of Iowa 375 Newton Road , 4156 MEBRF, Iowa City, IA 52242, USA Tel: 319 335 8260 Fax: 319 335 9565 Email: bento-soares@uiowa.edu The sequence contained an oligo-dT track that was present in the oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A tail. The sequence tag present in the cDNA between the NotI site and the oligo-dT track served to verify it as a clone from the non-normalized rat eye library cDNA Library Preparation: M.B. Soares Lab Clone Distribution: clones will be available through Research Genetics (www.resgen.com) Seq primer: M13 Forward POLYA=Yes.

```

ORIGIN
Query Match          76.8%; Score 19.2; DB 2; Length 391;
Best Local Similarity 87.5%; Pred. No. 9.6e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy  2  AAAAACACTTGGGAAGGGAGACTCA 25
    ||| ||| ||| ||| ||| ||| |||
Db  232 AAAGACACTTGCAGGGTGACTCA 209

RESULT 43
BI139768/c
LOCUS
DEFINITION
  BI139768          441 bp      mRNA          linear      EST 03-JUL-2001
  IPI_45 A05.b1_A002 Immature pannicle 1 (IPI) Sorghum bicolor cDNA,
  mRNA sequence.
ACCESSION
  BI139768
VERSION
  BI139768.1   GI:14592211
KEYWORDS
  EST.
SOURCE
  Sorghum bicolor (sorghum)
  Sorghum bicolor
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
  clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE
  1. (bases 1 to 441)
  Klein,R.R., Cordonnier-Pratt,M.-M., Gingle,A., Sudman,M. and
  Pratt,L.H.
  An EST database from Sorghum: developing preanthesis pannicles
  TITLE

```

Unpublished (2001)
 Contact: Cordomier-Pratt MM
 Laboratory for Genomics and Bioinformatics
 The University of Georgia, Department of Plant Biology
 Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
 Tel: 706 542 1860
 Fax: 706 583 0210
 Email: mmpratt@uga.edu
 Sequences have been trimmed to exclude PolyA, vector and regions
 below Phred quality 16. The threshold for high quality sequence is
 20. Three-prime sequences, which are obtained with PolyTMix or T7
 sequencing primer are presented as the reverse complement.
 Seq primer: JEN REV
 High quality sequence stop: 431
 POLYA=No.

FEATURES
 source
 1. .441
 /organism="Sorghum bicolor"
 /mol_type="RNA"
 /cultiivar="BTx623"
 /db_xref="taxon:4558"
 /clone_lib="Immature pannicle 1 (IPI)"
 /note="Organ: Developing preanthesis pannicles; Vector:
 pBluescript II SK(-) from Lambda Zap II; Site 1: XhoI;
 Site 2: SORI; The library was made from poly-A RNA in the
 cloning vector lambda ZAP II. Clones to be sequenced were
 prepared by mass excision."

ORIGIN
 Query Match 76.8%; Score 19.2; DB 2; Length 441;
 Best Local Similarity 87.5%; Pred. No. 9.8e+02;
 Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAAACACTTGGGAAGGAGACTCA 25
 |||||
 Db 363 AAAAAACACTTGC AAGGAGACTAA 340
 |||||

RESULT 44
 AQ228531/c
 LOCUS
 DEFINITION
 HS_2023_B2_G01_MR CIT Approved Human Genomic Sperm Library D Homo
 sapiens genomic clone Plate=2023 Col=2 Row=N, Genomic survey
 sequence.
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homidae; Homo.
 1 (bases 1 to 473)
 Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
 Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
 Hood,L.
 Sequence-tagged connectors: A sequence approach to mapping and
 scanning the human genome
 Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
 1049764
 Contact: Mahairas GG, Wallace JC, Hood L
 High Throughput Sequencing Center
 University of Washington
 401 Queen Anne Avenue North, Seattle, WA 98109, USA
 Tel: (206) 616-3618
 Fax: (206) 616-3887
 Email: jwallace@u.washington.edu
 Sequence Tagged Connector
 Plate: 2023 row: N column: 2
 Class: BAC ends
 High quality sequence stop: 473.

FEATURES
 source
 1. .473
 Location/Qualifiers


```

/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="plate=2023 Col=2 Row=N"
/sex="male"
/clone_lib="CIT Approved Human Genomic Sperm Library D"
/notes="Organ: sperm; Vector: pBelBAC11; BAC Clones in
E-Coli DH10B"

ORIGIN
Query Match          76.8%; Score 19.2; DB 9; Length 473;
Best Local Similarity 87.5%; Pred. No. 9.9e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAAGACTTGGAGGAGACTCA 25
    |||||
Db 90 AAGAACACTTGAAGGAGAGTCA 67

RESULT 45
CE564697
LOCUS
DEFINITION
tigr-gss-dog-1700032751217 Dog Library Canis familiaris genomic,
genomic survey sequence.
ACCESSION
CE564697
VERSION
CE564697.1 GI:36881478
KEYWORDS
GSS.
SOURCE
Canis familiaris (dog)
ORGANISM
Canis familiaris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;
Canis.
REFERENCE
1 (bases 1 to 495)
Kirksness,E.F., Balcer,V., Halpern,A.L., Levy,S., Remington,K.,
Rusch,D.B., DeCher,A.L., Pop,M., Wang,W., Fraser,C.M. and
Venter,J.C.
The dog genome: survey sequencing and comparative analysis
Science 301 (5641), 1998-1903 (2003)
14512627
Contact: Kirksness EF
The Institute for Genomic Research
Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
Rockville, MD 20850, USA
Tel: 301-838-0200
Fax: 301-838-0208
Email: ekirksnes@tigr.org
Class: shotgun.
FEATURES
Location/Qualifiers
source
1..495
/organism="Canis familiaris"
/mol_type="genomic DNA"
/strain="Standard Poodle"
/db_xref="taxon:9615"
/clone_lib="Dog Library"
/notes="Site 1: BetXi; Libraries were prepared from
peripheral blood"

ORIGIN
Query Match          76.8%; Score 19.2; DB 10; Length 495;
Best Local Similarity 87.5%; Pred. No. 9.9e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAAGACTTGGAGGAGACTCA 25
    |||||
Db 243 AAAAATACCTTGAAGAGAGGCTCA 266

RESULT 46
AA925291/c
LOCUS
DEFINITION
AA925291
UI-R-AI-ee-h-08-0-UI.81 UI-R-AI Rattus norvegicus cDNA clone
UI-R-AI-ee-h-08-0-UI 3', similar to gi|205029|gb|M11563|RATKALPS Rat
submaxillary gland PS kallikrein mRNA, complete cds, mRNA sequence.

```

```

ACCESSION
AA925291
VERSION
AA925291.1 GI:4236482
KEYWORDS
EST.
SOURCE
Rattus norvegicus (Norway rat)
ORGANISM
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE
1 (bases 1 to 498)
Bonaldo,M.F., Lennon,G. and Soares,M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
8889548
On Apr 21, 1998 this sequence version replaced gi:3072427.
Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road , 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
The sequence tag present in the cDNA between the NotI site and the
oligo-dT track served to identify it as a clone from the normalized
adult spleen library. cDNA library Preparation: M. Fatima Bonaldo,
Ph.D. Clone distribution: clones will be available through Research
Genetics This clone is also available through the I.M.A.G.E.
Consortium at LLNL (info@image.llnl.gov). IMAGE ID=1771334
Seq primer: M3 Forward
POLYA=NO.
FEATURES
Location/Qualifiers
source
1..498
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-AI-ee-h-08-0-UI"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="UI-R-AI"
/notes="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; The UI-R-AI
library is a subtracted library derived from the UI-R-AI
library. The UI-R-AI library consisted of a mixture of
individually tagged normalized libraries constructed from
rat placenta, adult lung, brain, liver, kidney, heart,
spleen, ovary, and muscle. The tag is a string of 3-5'
nucleotides present between the Not I site and the
oligo-dT track which allows identification of the library
of origin of a clone within the mixture. The subtracted
library (UI-R-AI) was constructed as follows: PCR
amplified cDNA inserts from a pool of approximately 3,840
UI-R-AI clones from which 3' ESTs had been derived was
used as a driver in a hybridization with the UI-R-AI
library in the form of single-stranded circles. The
remaining single-stranded circles (subtracted library) was
purified by hydroxyapatite column chromatography,
converted to double-stranded circles and electroporated
into DH10B bacteria (Life Technologies) to generate the
UI-R-AI library. This procedure has been previously
described (Bonaldo, Lennon and Soares, Genome Research 6:
791-806, 1996)"

```

ORIGIN

```

Query Match          76.8%; Score 19.2; DB 1; Length 498;
Best Local Similarity 87.5%; Pred. No. 9.9e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAAGACTTGGAGGAGACTCA 25
    |||||
Db 231 AAAGACACTTGAAGGAGTCA 208

```


Coordinated Laboratory for Computational Genomics

University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565

Email: bento-soares@uiowa.edu

The sequence contained an oligo-dT track that was present in the oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A tail. The sequence tag present in the cDNA between the NotI site and the oligo-dT track served to verify it as a clone from the non-normalized rat eye library cDNA Library Preparation: M.B. Soares Lab Clone distribution: clones will be available through Research Genetics (www.resgen.com)
Seq primer: M13 Forward
POLYA=Yes.

FEATURES

source Location/Qualifiers
1. .605

/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-DA0-byj-e-03-0-UI"
/dev_stage="ADULT"
/lab_host="DH10B (Life Technologies)"
/clone_lib="UI-R-DA0"
/notes="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; The UI-R-DA0 library is a non-normalized library constructed from rat salivary gland tissue. For a detailed description of the library from which this clone was derived, please visit our web site at ratest.eng.uiowa.edu. The subraction has been previously described in (Bonaldo, Lennon and Soares, Genome Research 6:791-806, 1996)
TAG_TISSUE=rat eye
TAG_LIB=UI-R-DA0
TAG_SEQ=CAGCC"

ORIGIN

Query Match 76.8%; Score 19.2; DB 2; Length 605;
Best Local Similarity 87.5%; Pred. No. 1e+03;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 AAAACACTTGGAGGGAGACTCA 25

Db 229 AAAGACACTTGCAGGGTGACTCA 206

RESULT 50

BI279445/c

LOCUS

DEFINITION BI279445 608 bp mRNA linear EST 19-JUL-2001
UI-R-DA0-by1-e-05-0-UI.s1 UI-R-DA0 Rattus norvegicus cDNA clone

UI-R-DA0-by1-e-05-0-UI 3', mRNA sequence.

ACCESSION BI279445

VERSION BI279445.1 GI:14927259

KEYWORDS EST.

SOURCE Rattus norvegicus (Norway rat)

ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE 1 (bases 1 to 608)

Bonaldo,M.F., Lennon,G. and Soares,M.B.

Normalization and subtraction: two approaches to facilitate gene discovery

Genome Res. 6 (9), 791-806 (1996)

JOURNAL 8889548

PUBMED

COMMENT

Contact: Soares, MB

Coordinated Laboratory for Computational Genomics

University of Iowa

375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA

Tel: 319 335 8250

Fax: 319 335 9565

Email: bento-soares@uiowa.edu

The sequence contained an oligo-dT track that was present in the oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A tail. The sequence tag present in the cDNA between the NotI site and the oligo-dT track served to verify it as a clone from the non-normalized rat eye library cDNA Library Preparation: M.B. Soares Lab Clone distribution: clones will be available through Research Genetics (www.resgen.com)
Seq primer: M13 Forward
POLYA=Yes.

FEATURES

source Location/Qualifiers
1. .608

/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-DA0-by1-e-05-0-UI"
/dev_stage="ADULT"
/lab_host="DH10B (Life Technologies)"
/clone_lib="UI-R-DA0"
/notes="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; The UI-R-DA0 library is a non-normalized library constructed from the salivary gland tissue. For a detailed description of the library from which this clone was derived, please visit our web site at ratest.eng.uiowa.edu. The subraction has been previously described in (Bonaldo, Lennon and Soares, Genome Research 6:791-806, 1996)
TAG_TISSUE=rat eye
TAG_LIB=UI-R-DA0
TAG_SEQ=CAGCC"

ORIGIN

Query Match 76.8%; Score 19.2; DB 2; Length 608;
Best Local Similarity 87.5%; Pred. No. 1e+03;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 AAAACACTTGGAGGGAGACTCA 25

Db 232 AAAGACACTTGCAGGGTGACTCA 209

Search completed: February 3, 2006, 22:02:04
Job time : 2960.67 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 3, 2006, 11:35:31 ; Search time 80.6667 Seconds
(without alignments)
550.897 Million cell updates/sec

Title: US-10-719-900-10

Perfect score: 25

Sequence: 1 aaaaacacttgaggagactca 25

-Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database :

Issued Patents NA:*

1: /cgn2_6/ptodata/1/ina/1 COMB.seq:*

2: /cgn2_6/ptodata/1/ina/5 COMB.seq:*

3: /cgn2_6/ptodata/1/ina/6A COMB.seq:*

4: /cgn2_6/ptodata/1/ina/6B COMB.seq:*

5: /cgn2_6/ptodata/1/ina/H COMB.seq:*

6: /cgn2_6/ptodata/1/ina/PTUS COMB.seq:*

7: /cgn2_6/ptodata/1/ina/PP COMB.seq:*

8: /cgn2_6/ptodata/1/ina/RE COMB.seq:*

9: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	18.2	72.8	601	3	US-09-949-016-28002
C 2	18.2	72.8	601	3	US-09-949-016-54169
C 3	18.2	72.8	601	3	US-09-949-016-142990
C 4	18.2	72.8	2092	3	US-09-023-655-936
C 5	18.2	72.8	5203	3	US-09-949-016-1562
C 6	18.2	72.8	6721	3	US-09-949-016-518
C 7	18.2	72.8	47858	3	US-09-949-016-14965
C 8	18.2	72.8	120213	3	US-09-949-016-13304
C 9	18.2	72.8	120217	3	US-09-949-016-12260
C 10	18.2	72.8	265038	3	US-09-949-016-15779
C 11	18.2	72.8	421118	3	US-09-949-016-16297
C 12	18.2	72.8	601	3	US-09-949-016-46494
C 13	17.8	71.2	601	3	US-09-949-016-46495
C 14	17.8	71.2	601	3	US-09-949-016-113510
C 15	17.8	71.2	601	3	US-09-949-016-113511
C 16	17.8	71.2	65744	3	US-09-949-016-12591
C 17	17.8	71.2	165998	3	US-09-949-016-15871
C 18	17.8	71.2	168734	3	US-09-949-016-15871
C 19	17.8	71.2	193689	3	US-09-949-016-12450
C 20	17.8	71.2	193689	3	US-09-949-016-12350
C 21	17.8	71.2	193689	3	US-09-949-016-13088
C 22	17.8	71.2	197496	3	US-09-877-177A-10
C 23	17.6	70.4	120	3	US-09-513-999C-17420
C 24	17.6	70.4	299	3	US-08-906-769-92

```

98 16.6 66.4 3 US-09-949-016-26792 Sequence 26792, A
99 16.6 66.4 3 US-09-949-016-32735 Sequence 32735, A
100 16.6 66.4 3 US-09-949-016-32736 Sequence 32736, A
101 16.6 66.4 3 US-09-949-016-42795 Sequence 42795, A
102 16.6 66.4 3 US-09-949-016-42796 Sequence 42796, A
103 16.6 66.4 3 US-09-949-016-42796 Sequence 42796, A
104 16.6 66.4 3 US-09-949-016-50585 Sequence 50585, A
105 16.6 66.4 3 US-09-949-016-50585 Sequence 50585, A
106 16.6 66.4 3 US-09-949-016-69121 Sequence 69121, A
107 16.6 66.4 3 US-09-949-016-69121 Sequence 69121, A
108 16.6 66.4 3 US-09-949-016-142348 Sequence 142348, A
109 16.6 66.4 3 US-09-949-016-142349 Sequence 142349, A
110 16.6 66.4 3 US-09-949-016-764 Sequence 764, A
111 16.6 66.4 3 US-09-949-016-764 Sequence 764, A
112 16.6 66.4 3 US-09-949-016-4025 Sequence 4025, A
113 16.6 66.4 3 US-09-949-016-4025 Sequence 4025, A
114 16.6 66.4 3 US-09-949-016-4025 Sequence 4025, A
115 16.6 66.4 3 US-09-949-016-4025 Sequence 4025, A
116 16.6 66.4 3 US-09-949-016-4025 Sequence 4025, A
117 16.6 66.4 3 US-09-949-016-4025 Sequence 4025, A
118 16.6 66.4 3 US-09-949-016-4025 Sequence 4025, A
119 16.6 66.4 3 US-09-949-016-4025 Sequence 4025, A
120 16.6 66.4 3 US-09-949-016-4025 Sequence 4025, A
121 16.6 66.4 3 US-09-949-016-4025 Sequence 4025, A
122 16.6 66.4 3 US-09-949-016-4025 Sequence 4025, A
123 16.6 66.4 3 US-09-949-016-4025 Sequence 4025, A
124 16.6 66.4 3 US-09-949-016-4025 Sequence 4025, A
125 16.6 66.4 3 US-09-949-016-4025 Sequence 4025, A
126 16.6 66.4 3 US-09-949-016-4025 Sequence 4025, A
127 16.6 66.4 3 US-09-949-016-4025 Sequence 4025, A
128 16.6 66.4 3 US-09-949-016-4025 Sequence 4025, A
129 16.6 66.4 3 US-09-949-016-4025 Sequence 4025, A
130 16.6 66.4 3 US-09-949-016-4025 Sequence 4025, A
131 16.6 66.4 3 US-09-949-016-4025 Sequence 4025, A
132 16.6 66.4 3 US-09-949-016-4025 Sequence 4025, A
133 16.6 66.4 3 US-09-949-016-4025 Sequence 4025, A
134 16.6 66.4 3 US-09-949-016-4025 Sequence 4025, A
135 16.6 66.4 3 US-09-949-016-4025 Sequence 4025, A
136 16.6 66.4 3 US-09-949-016-4025 Sequence 4025, A
137 16.6 66.4 3 US-09-949-016-4025 Sequence 4025, A
138 16.6 66.4 3 US-09-949-016-4025 Sequence 4025, A
139 16.6 66.4 3 US-09-949-016-4025 Sequence 4025, A
140 16.6 66.4 3 US-09-949-016-4025 Sequence 4025, A
141 16.6 66.4 3 US-09-949-016-4025 Sequence 4025, A
142 16.6 66.4 3 US-09-949-016-4025 Sequence 4025, A
143 16.6 66.4 3 US-09-949-016-4025 Sequence 4025, A
144 16.6 66.4 3 US-09-949-016-4025 Sequence 4025, A
145 16.6 66.4 3 US-09-949-016-4025 Sequence 4025, A
146 16.6 66.4 3 US-09-949-016-4025 Sequence 4025, A
147 16.6 66.4 3 US-09-949-016-4025 Sequence 4025, A
148 16.6 66.4 3 US-09-949-016-4025 Sequence 4025, A
149 16.6 66.4 3 US-09-949-016-4025 Sequence 4025, A
150 16.6 66.4 3 US-09-949-016-4025 Sequence 4025, A
```

ALIGNMENTS

```

RESULT 1
US-09-949-016-28002/c
; Sequence 28002, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 142990

US-09-949-016-142990
; Sequence 142990, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 142990
```

```
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-142990

Query Match          72.8%; Score 18.2; DB 3; Length 601;
Best Local Similarity 87.0%; Pred. No. 48;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAACACTTGGAGGAGACTC 24
Db 530 AAAAACAACATGGAAGGACAGTC 552

RESULT 4
US-09-023-655-936
; Sequence 936, Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; TITLE OF INVENTION: COMPOSITION: EXPRESSION
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/023,655
; FILING DATE: HERewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 853-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 936:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2092 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: GI381145
US-09-023-655-936

Query Match          72.8%; Score 18.2; DB 3; Length 2092;
Best Local Similarity 87.0%; Pred. No. 60;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAACAACCTTGGAGGAGACT 23
Db 724 AAAACAACACTTGGCAAGGAGACT 746

RESULT 5
US-09-949-016-1562
; Sequence 1562, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1562
; LENGTH: 5203
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-1562

Query Match          72.8%; Score 18.2; DB 3; Length 5203;
Best Local Similarity 87.0%; Pred. No. 72;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAACAACCTTGGAGGAGACT 23
Db 1613 AAAACAACACTTGGCAAGGAGACT 1635

RESULT 6
US-09-949-016-518
; Sequence 518, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 518
; LENGTH: 6721
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-518

Query Match          72.8%; Score 18.2; DB 3; Length 6721;
Best Local Similarity 87.0%; Pred. No. 75;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAACAACCTTGGAGGAGACT 23
Db 1618 AAAACAACACTTGGCAAGGAGACT 1640

RESULT 7
US-09-949-016-14965/c
; Sequence 14965, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
```

```

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14965
; LENGTH: 47858
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(47858)
; OTHER INFORMATION: n = A,T,C or G
; US-09-949-016-14965

```

Query Match	72.8%	Score 18.2;	DB 3;	Length 47858;
Best Local Similarity	87.0%	Pred. NO. 1.1e+02;		
Matches 20:	Conservative	0;	Mismatches 3;	Indels 0;
				Gaps 0;

Oy

1 AAAAAACACTTGGAAAGGGACT 23
||| |||||||||
D6

7963 AACAAAACACTTGGAAAGGCACT 7941

```

RESULT 8
US-09-949-016-13304
; Sequence 13304, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13304
; LENGTH: 120213
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(120213)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13304

```

Query Match	72.8%;	Score 18.2;	DB 3;	Length 120213;
Best Local Similarity	87.0%;	Pred. No. 1.3e+02;		
Matches	20:	Conservative	0: Mismatches	3: Indels
			0: Gaps	0:

Qy 1 AAAAAACACTTGGAAAGGAGACT 23
||| ||| ||| ||| ||| ||| ||| |||
db 84632 AAACAACACTTGGCAAGGAGACT 84654

RESULT 9
US-09-949-016-12260
: Sequence 12260, Application US/09949016

```

; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CU001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12260
; LENGTH: 120217
; TYPE: DNA
; ORGANISM: Human
; FEATURES:
; NAME/KEY: misc_feature
; LOCATION: (1)..(120217)
; OTHER INFORMATION: n = A,T,C or G
; JS-09-949-016-12260

```

Query Match	72.8%	Score 18.2;	DB 3;	Length 120217;
Best Local Similarity	87.0%;	Pred. No. 1.3e+02;		
Matches	20: Conservative	0: Mismatches	3: Indels	0: Gaps

Qy

1 AAAAAACACTTGGAAAGGAGCT 23
||| ||| ||| ||| ||| ||| ||| |||
Dd

84636 AAACAACACTTGGCAAGGAGCT 84658

```

RESULT 10
US-09-949-016-15779/c
; Sequence 15779, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15779
; LENGTH: 265038
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(265038)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15779

```

Query Match	72.8%;	Score 18.2;	DB 3;	Length 265038;
Best Local Similarity	87.0%;	Pred. No. 1.5e+02;		
Matches 20;	Conservative	0;	Mismatches 3;	Indels 0;
Gaps	0;			

Qy 2 AAAACACTTGAAGGGAGCTC 24
 |||||
 Db 163302 AAAACACATGGAAGGGACAGTC 163280


```
RESULT 11
US-09-949-016-16297/c
; Sequence 16297, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16297
; LENGTH: 421118
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(421118)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16297

Query Match 72.8%; Score 18.2; DB 3; Length 421118;
Best Local Similarity 87.0%; Pred. No. 1.6e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAACACTTGGGAAGGAGACT 23
Db 59077 AAACACTACTTGGGAAGGAGAAAT 59055

RESULT 12
US-09-949-016-46494/c
; Sequence 46494, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 46494
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-46494

Query Match 71.2%; Score 17.8; DB 3; Length 601;
Best Local Similarity 90.5%; Pred. No. 74;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAACACTTGGGAAGGAGGA 21
Db 539 ATAAACACTTAGGAAGGAGGA 519

RESULT 13
US-09-949-016-113511/c
; Sequence 113511, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 113511
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-113511

Query Match 71.2%; Score 17.8; DB 3; Length 601;
Best Local Similarity 90.5%; Pred. No. 74;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAACACTTGGGAAGGAGGA 21
Db 539 ATAAACACTTAGGAAGGAGGA 519

RESULT 14
US-09-949-016-113510/c
; Sequence 113510, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 113510
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-113510

Query Match 71.2%; Score 17.8; DB 3; Length 601;
Best Local Similarity 90.5%; Pred. No. 74;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAACACTTGGGAAGGAGGA 21
Db 379 ATAAACACTTAGGAAGGAGGA 359

RESULT 15
US-09-949-016-113511/c
; Sequence 113511, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
```

```

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 113511
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
;
US-09-949-016-113511

Query Match 71.2%; Score 17.8; DB 3; Length 601;
Best Local Similarity 90.5%; Pred. No. 74;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAACACTTGGGAAGGGAGA 21
| | | | | | | | | | | | | | | |
Db 379 ATAAAAACACTAGGAAGGGAGA 359

RESULT 16
US-09-949-016-12591
; Sequence 12591, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12591
; LENGTH: 65744
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(65744)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12591

Query Match 71.2%; Score 17.8; DB 3; Length 65744;
Best Local Similarity 90.5%; Pred. No. 1.8e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAACACTTGGGAAGGGAGA 21
| | | | | | | | | | | | | | | |
Db 34052 AAAAAACAGTTGAAGGAAGA 34072

RESULT 17
US-09-949-016-15871
; Sequence 15871, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

```

```
/ NAME/KEY: exon
/ LOCATION: (136117)...(136261)
/ NAME/KEY: intron
/ LOCATION: (136262)...(137936)
/ NAME/KEY: exon
/ LOCATION: (137937)...(138053)
/ NAME/KEY: intron
/ LOCATION: (138054)...(138637)
/ NAME/KEY: exon
/ LOCATION: (138638)...(138766)
/ NAME/KEY: intron
/ LOCATION: (138767)...(138864)
/ NAME/KEY: exon
/ LOCATION: (138865)...(138940)
/ NAME/KEY: intron
/ LOCATION: (138941)...(139765)
/ NAME/KEY: exon
/ LOCATION: (139766)...(139860)
/ NAME/KEY: intron
/ LOCATION: (139861)...(142245)
/ NAME/KEY: exon
/ LOCATION: (142246)...(142445)
/ NAME/KEY: intron
/ LOCATION: (142446)...(143605)
/ NAME/KEY: exon
/ LOCATION: (143606)...(143738)
/ NAME/KEY: intron
/ LOCATION: (143739)...(145838)
/ NAME/KEY: exon
/ LOCATION: (145839)...(145931)
/ NAME/KEY: intron
/ LOCATION: (145932)...(147385)
/ NAME/KEY: exon
/ LOCATION: (147386)...(147544)
/ NAME/KEY: intron
/ LOCATION: (147545)...(153274)
/ NAME/KEY: exon
/ LOCATION: (153275)...(153321)
/ NAME/KEY: intron
/ LOCATION: (153322)...(155088)
/ NAME/KEY: exon
/ LOCATION: (155089)...(155231)
/ NAME/KEY: intron
/ LOCATION: (155232)...(156025)
/ NAME/KEY: exon
/ LOCATION: (156026)...(156151)
/ NAME/KEY: intron
/ LOCATION: (156152)...(156826)
/ NAME/KEY: exon
/ LOCATION: (156827)...(156928)
/ NAME/KEY: intron
/ LOCATION: (156929)...(163399)
/ NAME/KEY: exon
/ LOCATION: (163400)...(163586)
/ NAME/KEY: intron
/ LOCATION: (163587)...(163586)
US-09-676-610B-24
```

```
Query Match 71.2%; Score 17.8; DB 3; Length 169998;
Best Local Similarity 90.5%; Pred. No. 2.1e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy 2 AAAAACAACCTTGGAGGGGAGAC 22
Db 104105 AAAAACAACCTTGGAGGGGAGAC 104125
```

```
RESULT 19
US-09-949-016-14870/c
; Sequence 14870, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
```

```
/ FILE REFERENCE: CL001307
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ CURRENT FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 14870
/ LENGTH: 186734
/ TYPE: DNA
/ ORGANISM: Human
US-09-949-016-14870
```

```
Query Match 71.2%; Score 17.8; DB 3; Length 186734;
Best Local Similarity 90.5%; Pred. No. 2.2e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy 1 AAAAACAACCTTGGAGGGGAGA 21
Db 23486 ATAAACAACCTAGGAGGGAGA 23466
```

```
RESULT 20
US-09-949-016-12350/c
; Sequence 12350, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12350
; LENGTH: 193689
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12350
```

```
Query Match 71.2%; Score 17.8; DB 3; Length 193689;
Best Local Similarity 90.5%; Pred. No. 2.2e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy 1 AAAAACAACCTTGGAGGGGAGA 21
Db 23486 ATAAACAACCTAGGAGGGAGA 23466
```

```
RESULT 21
US-09-949-016-13088/c
; Sequence 13088, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
```

; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13088
; LENGTH: 193689
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13088

Query Match 71.2%; Score 17.8; DB 3; Length 193689;
Best Local Similarity 90.5%; Pred. No. 2.2e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 AAAAAACACTTGGAGGGGAGA 21
Db 23486 ATAAACACTAGGAGGGGAGA 23466

RESULT 22
US-09-877-177A-10
; Sequence 10, Application US/09877177A
; Patent No. 6582919
; GENERAL INFORMATION:
; APPLICANT: K. Danenberg
; TITLE OF INVENTION: Method of determining Epidermal Growth
; TITLE OF INVENTION: Factor Receptor and HER2-Neu Gene Expression
; TITLE OF INVENTION: and Correlation of Levels Thereof With Survival
; FILE REFERENCE: 11220/120
; CURRENT APPLICATION NUMBER: US/09/877,177A
; CURRENT FILING DATE: 2001-06-11
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 197496
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-877-177A-10

Query Match 71.2%; Score 17.8; DB 3; Length 197496;
Best Local Similarity 90.5%; Pred. No. 2.2e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 2 AAAAAACACTTGGAGGGGAGAC 22
Db 112105 AAAAGACTTGGAGGGGAGAC 112125

RESULT 23
US-09-513-999C-17420/c
; Sequence 17420, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 17420
; LENGTH: 120
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-513-999C-17420

Query Match 70.4%; Score 17.6; DB 3; Length 120;
Best Local Similarity 83.3%; Pred. No. 67;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 AAAAAACACTTGGAGGGGAGACTC 24
Db 114 AAAAAAAGCTTGAAGTTAGACTC 91

RESULT 24
US-08-906-769-92
; Sequence 92, Application US/08906769
; Patent No. 6077687
; GENERAL INFORMATION:
; APPLICANT: Grieve, Robert B.
; APPLICANT: Rushlow, Keith E.
; APPLICANT: Wu Hunter, Shirley
; APPLICANT: Frank, Glenn R.
; APPLICANT: Stiegler, Gary
; APPLICANT: Gaines, Patrick J.
; APPLICANT: Silver, Gary
; TITLE OF INVENTION: FLEA PROTEASE PROTEINS, NUCLEIC ACID
; TITLE OF INVENTION: MOLECULES AND USES THEREOF
; NUMBER OF SEQUENCES: 190
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheridan Ross & McIntosh
; STREET: 1700 Lincoln Street, Suite 3500
; CITY: Denver
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80203
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/906,769
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/639,075
; FILING DATE: 24-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Connell, Gary J.
; REGISTRATION NUMBER: 32,020
; REFERENCE/DOCKET NUMBER: 2618-25-C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 863-9700
; TELEFAX: (303) 863-0223
; INFORMATION FOR SEQ ID NO: 92:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 299 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 3..299
; OTHER INFORMATION: /note= "At pos. bp 178/179, change
; OTHER INFORMATION: G/C to V/Y. At pos. aa 59, substitue Xaa."
US-08-906-769-92

Query Match 70.4%; Score 17.6; DB 3; Length 299;
Best Local Similarity 83.3%; Pred. No. 80;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 2 AAAAAACACTTGGAGGGGAGACTCA 25
Db 267 AAAGACACTTGCAGGGAGATTCA 290

RESULT 25
US-08-906-616-92
; Sequence 92, Application US/08906616
; Patent No. 6121035
; GENERAL INFORMATION:
; APPLICANT: Grieve, Robert B.
; APPLICANT: Rushlow, Keith E.
; APPLICANT: Wu Hunter, Shirley
; APPLICANT: Frank, Glenn R.
; APPLICANT: Stiegler, Gary
; APPLICANT: Gaines, Patrick J.
; APPLICANT: Silver, Gary
; TITLE OF INVENTION: FLEA AMINOPEPTIDASE PROTEINS AND USES THEREOF
; NUMBER OF SEQUENCES: 190
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheridan Ross P.C.
; STREET: 1700 Lincoln Street, Suite 3500
; CITY: Denver
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80203
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/906,616
; FILING DATE: 05-AUG-1997
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Connell, Gary J.
; REGISTRATION NUMBER: 32,020
; REFERENCE/DOCKET NUMBER: 2618-25-C2-3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 863-9700
; TELEFAX: (303) 863-0223
; INFORMATION FOR SEQ ID NO: 92:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 299 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 3..299
; OTHER INFORMATION: /note= "At pos. bp 178/179, change
; G/C to V/Y. At pos. aa 59, substitute Xaa."
US-08-906-616-92

Query Match 70.4%; Score 17.6; DB 3; Length 299;
Best Local Similarity 83.3%; Pred. No. 80;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AAAAACTTGGAGGGAGACTCA 25
||| ||||| ||||| ||||| |||||
Db 267 AAAGACACTTGGCCAGGGAGATTCA 290

RESULT 26
US-08-817-795-92
; Sequence 92, Application US/08817795
; Patent No. 6139840
; GENERAL INFORMATION:
; APPLICANT: Grieve, Robert B.
; APPLICANT: Rushlow, Keith E.
; APPLICANT: Hunter, Shirley Wu
; APPLICANT: Frank, Glenn R.
; APPLICANT: Heath, Andrew W.
; APPLICANT: Yamaka, Miles Yamanaka
; APPLICANT: Arfsten, Ann
; APPLICANT: Dale, Beverly

; APPLICANT: Stiegler, Gary
; TITLE OF INVENTION: USE OF PROTEASE INHIBITORS AND
; PROTEASE VACCINES TO PROTECT ANIMALS FROM FLEA
; INFESTATION, AND FLEA PROTEASE PROTEINS, NUCLEIC ACID
; MOLECULES, AND USES THEREOF
; NUMBER OF SEQUENCES: 119
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheridan Ross & McIntosh
; STREET: 1700 Lincoln Street, Suite 3500
; CITY: Denver
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80203
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/817,795
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/14442
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Gary J. Connell
; REGISTRATION NUMBER: 32,020
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 863-9700
; TELEFAX: (303) 863-0223
; INFORMATION FOR SEQ ID NO: 92:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 299 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 3..299
; NAME/KEY: Xaa = any amino acid
; LOCATION: 59
US-08-817-795-92

Query Match 70.4%; Score 17.6; DB 3; Length 299;
Best Local Similarity 83.3%; Pred. No. 80;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AAAAACTTGGAGGGAGACTCA 25
||| ||||| ||||| ||||| |||||
Db 267 AAAGACACTTGGCCAGGGAGATTCA 290

RESULT 27
US-08-639-075A-92
; Sequence 92, Application US/08639075A
; Patent No. 6150125
; GENERAL INFORMATION:
; APPLICANT: Grieve, Robert B.
; APPLICANT: Rushlow, Keith E.
; APPLICANT: Wu Hunter, Shirley
; APPLICANT: Frank, Glenn R.
; APPLICANT: Stiegler, Gary
; APPLICANT: Gaines, Patrick J.
; APPLICANT: Silver, Gary
; TITLE OF INVENTION: FLEA PROTEASE PROTEINS, NUCLEIC ACID
; MOLECULES AND USES THEREOF
; NUMBER OF SEQUENCES: 190
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheridan Ross & McIntosh

STREET: 1700 Lincoln Street, Suite 3500
CITY: Denver
STATE: Colorado
COUNTRY: USA
ZIP: 80203
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/639,075A
FILING DATE: 24-APR-1996
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Connell, Gary J.
REGISTRATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 2618-25-C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO: 92:
SEQUENCE CHARACTERISTICS:
LENGTH: 299 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 3..299
OTHER INFORMATION: /note= "At pos. bp 178/179, change
G/C to V/Y. At pos. aa 59, substitute Xaa."
US-08-639-075A.92

Query Match : 70.4%; Score 17.6; DB 3; Length 299;
Best Local Similarity 83.3%; Pred. No. 80;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AAAAACTTGAAGGAGACTCA 25
Db 267 AAGACACTTGCAGGAGATTCA 290

RESULT 28

US-09-012-431-92
Sequence 92, Application US/09012431
Patent No. 6180383

GENERAL INFORMATION:

APPLICANT: Grieve, Robert B.
Rushlow, Keith E.
Wu Hunter, Shirley
Frank, Glenn R.
Stiegler, Gary
Gaines, Patrick J.
Silver, Gary

TITLE OF INVENTION: FLEA PROTEASE PROTEINS, NUCLEIC ACID
MOLECULES AND USES THEREOF

NUMBER OF SEQUENCES: 190

CORRESPONDENCE ADDRESS:

ADDRESSEE: Sheridan Ross & McIntosh

STREET: 1700 Lincoln Street, Suite 3500

CITY: Denver

STATE: Colorado

COUNTRY: USA

ZIP: 80203

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/012,431

FILING DATE: 23-Jan-1998
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/639,075
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Connell, Gary J.
REGISTRATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 2618-25-C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO: 92:
SEQUENCE CHARACTERISTICS:
LENGTH: 299 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 3..299
OTHER INFORMATION: /note= "At pos. bp 178/179, change
G/C to V/Y. At pos. aa 59, substitute Xaa."
US-09-012-431-92

Query Match : 70.4%; Score 17.6; DB 3; Length 299;
Best Local Similarity 83.3%; Pred. No. 80;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AAAAACTTGAAGGAGACTCA 25

Db 267 AAGACACTTGCAGGAGATTCA 290

RESULT 29

US-09-012-692-92

Sequence 92, Application US/09012692

Patent No. 6214579

GENERAL INFORMATION:

APPLICANT: Grieve, Robert B.
Rushlow, Keith E.
Wu Hunter, Shirley
Frank, Glenn R.
Stiegler, Gary
Gaines, Patrick J.

APPLICANT: Silver, Gary

TITLE OF INVENTION: FLEA PROTEASE PROTEINS, NUCLEIC ACID

MOLECULES AND USES THEREOF

NUMBER OF SEQUENCES: 190

CORRESPONDENCE ADDRESS:

ADDRESSEE: Sheridan Ross & McIntosh

STREET: 1700 Lincoln Street, Suite 3500

CITY: Denver

STATE: Colorado

COUNTRY: USA

ZIP: 80203

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/012,692

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/639,075

FILING DATE: 24-APR-1996

ATTORNEY/AGENT INFORMATION:

NAME: Connell, Gary J.

REGISTRATION NUMBER: 32,020

```

; REFERENCE/DOCKET NUMBER: 2618-25-C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 863-9700
; TELEFAX: (303) 863-0223
; INFORMATION FOR SEQ ID NO: 92:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 299 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 3..299
; OTHER INFORMATION: /note= "At pos. bp 178/179, change
; OTHER INFORMATION: G/C to V/Y. At pos. aa 59, substitute Xaa."
US-09-012-692-92

Query Match 70.4%; Score 17.6; DB 3; Length 299;
Best Local Similarity 83.3%; Pred. No. 80;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AAAAAGACTTGGAGGAGACTCA 25
||| ||||| ||||| ||||| |||||
DB 267 AAAGACACTTGCAGGAGATTCA 290

RESULT 30
US-08-906-613-92
; Sequence 92, Application US/08906613
; Patent No. 6232096
; GENERAL INFORMATION:
; APPLICANT: Grieve, Robert B.
; APPLICANT: Rushlow, Keith E.
; APPLICANT: Wu Hunter, Shirley
; APPLICANT: Frank, Glenn R.
; APPLICANT: Stiegler, Gary
; APPLICANT: Gaines, Patrick J.
; APPLICANT: Silver, Gary
; TITLE OF INVENTION: FLEA PROTEASE PROTEINS, NUCLEIC ACID
; TITLE OF INVENTION: MOLECULES AND USES THEREOF
; NUMBER OF SEQUENCES: 190
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheridan Ross & McIntosh
; STREET: 1700 Lincoln Street, Suite 3500
; CITY: Denver
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80203
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/906,613
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/639,075
; FILING DATE: 24-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Connell, Gary J.
; REGISTRATION NUMBER: 32,020
; REFERENCE/DOCKET NUMBER: 2618-25-C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 863-9700
; TELEFAX: (303) 863-0223
; INFORMATION FOR SEQ ID NO: 92:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 299 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 3..299
; OTHER INFORMATION: /note= "At pos. bp 178/179, change
; OTHER INFORMATION: G/C to V/Y. At pos. aa 59, substitute Xaa."
US-09-012-692-92

Query Match 70.4%; Score 17.6; DB 3; Length 299;
Best Local Similarity 83.3%; Pred. No. 80;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AAAAAGACTTGGAGGAGACTCA 25
||| ||||| ||||| ||||| |||||
DB 267 AAAGACACTTGCAGGAGATTCA 290

RESULT 31
PCT-US95-14442A-92
; Sequence 92, Application PC/TUS9514442A
; GENERAL INFORMATION:
; APPLICANT: Grieve, Robert B.
; APPLICANT: Rushlow, Keith E.
; APPLICANT: Hunter, Shirley Wu
; APPLICANT: Frank, Glenn R.
; APPLICANT: Heath, Andrew W.
; APPLICANT: Yamaka, Miles Yamanaka
; APPLICANT: Arfsten, Ann
; APPLICANT: Dale, Beverly
; APPLICANT: Stiegler, Gary
; TITLE OF INVENTION: USE OF PROTEASE INHIBITORS AND
; TITLE OF INVENTION: PROTEASE VACCINES TO PROTECT ANIMALS FROM FLEA
; TITLE OF INVENTION: INFESTATION, AND FLEA PROTEASE PROTEINS, NUCLEIC ACID
; TITLE OF INVENTION: MOLECULES, AND USES THEREOF
; NUMBER OF SEQUENCES: 119
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheridan Ross & McIntosh
; STREET: 1700 Lincoln Street, Suite 3500
; CITY: Denver
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80203
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/14442A
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Gary J. Connell
; REGISTRATION NUMBER: 32,020
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 863-9700
; TELEFAX: (303) 863-0223
; INFORMATION FOR SEQ ID NO: 92:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 299 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 3..299
; OTHER INFORMATION: Xaa = any amino acid
; NAME/KEY: Xaa = any amino acid
; LOCATION: 59
; PCT-US95-14442A-92

```

Query Match 70.4%; Score 17.6; DB 6; Length 299;
Best Local Similarity 83.3%; Pred. No. 80;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AAAAACACTTGAAGGAGACTCA 25
|||||
Db 267 AAAGACACTTGCAGGAGATTCA 290
|||||

RESULT 32
US-08-906-769-122
; Sequence 122, Application US/08906769
; Patent No. 6077687
; GENERAL INFORMATION:
; APPLICANT: Grieve, Robert B.
; APPLICANT: Rushlow, Keith E.
; APPLICANT: Wu Hunter, Shirley
; APPLICANT: Frank, Glenn R.
; APPLICANT: Stiegler, Gary
; APPLICANT: Gaines, Patrick J.
; APPLICANT: Silver, Gary
; TITLE OF INVENTION: FLEA PROTEASE PROTEINS, NUCLEIC ACID
; NUMBER OF SEQUENCES: 190
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheridan Ross & McIntosh
; STREET: 1700 Lincoln Street, Suite 3500
; CITY: Denver
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80203
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/906,769
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/639,075
; FILING DATE: 24-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Connell, Gary J.
; REGISTRATION NUMBER: 32,020
; REFERENCE/DOCKET NUMBER: 2618-25-C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 863-9700
; TELEFAX: (303) 863-0223
; INFORMATION FOR SEQ ID NO: 122:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 436 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2..436
; OTHER INFORMATION: /note= "at pos. bp 301, change A to
; OTHER INFORMATION: W; at pos. bp 342, change C to Y; at pos. bp 397, change C to
; OTHER INFORMATION: at pos. bp 431, change G to S. At pos. aa 100, 114 and 144,
; OTHER INFORMATION: substitute xaa."
US-08-906-769-122

Query Match 70.4%; Score 17.6; DB 3; Length 436;
Best Local Similarity 83.3%; Pred. No. 86;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AAAAACACTTGAAGGAGACTCA 25
|||||

Db 254 AAAGACACTTGCAGGAGATTCA 277

RESULT 33
US-08-906-616-122
; Sequence 122, Application US/08906616
; Patent No. 6121035
; GENERAL INFORMATION:
; APPLICANT: Grieve, Robert B.
; APPLICANT: Rushlow, Keith E.
; APPLICANT: Wu Hunter, Shirley
; APPLICANT: Frank, Glenn R.
; APPLICANT: Stiegler, Gary
; APPLICANT: Gaines, Patrick J.
; APPLICANT: Silver, Gary
; TITLE OF INVENTION: FLEA AMINOPEPTIDASE PROTEINS AND USES THEREOF
; NUMBER OF SEQUENCES: 190
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheridan Ross P.C.
; STREET: 1700 Lincoln Street, Suite 3500
; CITY: Denver
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80203
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/906,616
; FILING DATE: 05-AUG-1997
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Connell, Gary J.
; REGISTRATION NUMBER: 32,020
; REFERENCE/DOCKET NUMBER: 2618-25-C2-3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 863-9700
; TELEFAX: (303) 863-0223
; INFORMATION FOR SEQ ID NO: 122:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 436 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2..436
; OTHER INFORMATION: /note= "at pos. bp 301, change A to
; OTHER INFORMATION: W; at pos. bp 342, change C to Y; at pos. bp 397, change C to
; OTHER INFORMATION: at pos. bp 431, change G to S. At pos. aa 100, 114 and 144,
; OTHER INFORMATION: substitute xaa."
US-08-906-616-122

Query Match 70.4%; Score 17.6; DB 3; Length 436;
Best Local Similarity 83.3%; Pred. No. 86;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AAAAACACTTGAAGGAGACTCA 25
|||||

Db 254 AAAGACACTTGCAGGAGATTCA 277
|||||

RESULT 34
US-08-639-075A-122
; Sequence 122, Application US/08639075A
; Patent No. 6150125
; GENERAL INFORMATION:
; APPLICANT: Grieve, Robert B.
; APPLICANT: Rushlow, Keith E.
; APPLICANT: Wu Hunter, Shirley

APPLICANT: Frank, Glenn R.
APPLICANT: Stiegler, Gary
APPLICANT: Gaines, Patrick J.
APPLICANT: Silver, Gary
TITLE OF INVENTION: FLEA PROTEASE PROTEINS, NUCLEIC ACID
AND USES THEREOF
NUMBER OF SEQUENCES: 190
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sheridan Ross & McIntosh
STREET: 1700 Lincoln Street, Suite 3500
CITY: Denver
STATE: Colorado
COUNTRY: USA
ZIP: 80203
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/639,075A
FILING DATE: 24-APR-1996
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Connell, Gary J.
REGISTRATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 2618-25-C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223
SEQUENCE CHARACTERISTICS:
LENGTH: 436 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 2..436
OTHER INFORMATION: /note= "At pos. bp 301, change A to
W; at pos. bp 342, change C to Y; at pos. bp 397, change C to
G; at pos. bp 431, change G to S. At pos. aa 100, 114 and 144,
substitute Xaa."
US-08-639-075A-122
Query Match 70.4%; Score 17.6; DB 3; Length 436;
Best Local Similarity 83.3%; Pred. No. 86;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Qy 2 AAAAACACTTGGAGGGAGACTCA 25
Db 254 AAAGACACTTGGCCAGGGAGATTCA 277
RESULT 35
US-09-012-431-122
Sequence 122, Application US/09012431
Patent No. 6180383
GENERAL INFORMATION:
APPLICANT: Grieve, Robert B.
Rushlow, Keith E.
Wu Hunter, Shirley
Frank, Glenn R.
Stiegler, Gary
Gaines, Patrick J.
Silver, Gary
TITLE OF INVENTION: FLEA PROTEASE PROTEINS, NUCLEIC ACID
AND USES THEREOF
NUMBER OF SEQUENCES: 190
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sheridan Ross & McIntosh
STREET: 1700 Lincoln Street, Suite 3500

CITY: Denver
STATE: Colorado
COUNTRY: USA
ZIP: 80203
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/012,431
FILING DATE: 23-Jan-1998
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/639,075
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Connell, Gary J.
REGISTRATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 2618-25-C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223
SEQUENCE CHARACTERISTICS:
LENGTH: 436 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 2..436
OTHER INFORMATION: /note= "At pos. bp 301, change A to
W; at pos. bp 342, change C to Y; at pos. bp 397, change C to Y;
at pos. bp 431, change G to S. At pos. aa 100, 114 and 144,
substitute Xaa."
US-09-012-431-122
Query Match 70.4%; Score 17.6; DB 3; Length 436;
Best Local Similarity 83.3%; Pred. No. 86;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Qy 2 AAAAACACTTGGAGGGAGACTCA 25
Db 254 AAAGACACTTGGCCAGGGAGATTCA 277
RESULT 36
US-09-032-215-19
Sequence 19, Application US/09032215
Patent No. 6204010
GENERAL INFORMATION:
APPLICANT: Stiegler, Gary L.
Gaines, Patrick J.
TITLE OF INVENTION: FLEA PROTEASE PROTEINS, NUCLEIC
ACID MOLECULES, AND USES THEREOF
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sheridan Ross P.C.
STREET: 1700 Lincoln Street, Suite 3500
CITY: Denver
STATE: Colorado
COUNTRY: U.S.A.
ZIP: 80203
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII DOS TEXT
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/032,215

;
; FILING DATE: 27-FEB-1998
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Connell, Gary J.
; REGISTRATION NUMBER: 32,020
; REFERENCE/DOCKET NUMBER: 2618-25-C6
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 863-9700
; TELEFAX: (303) 863-0223
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 436 nucleotides
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..436
; FEATURE:
; NAME/KEY: W = A
; LOCATION: 301
; FEATURE:
; NAME/KEY: Y = C
; LOCATION: 342, 397
; FEATURE:
; NAME/KEY: S = C
; LOCATION: 431
; FEATURE:
; NAME/KEY: Xaa = Unknown
; LOCATION: 100, 114, 144
; US-09-032-215-19

Query Match 70.4%; Score 17.6; DB 3; Length 436;
Best Local Similarity 83.3%; Pred. No. 86;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AAAACACTTGAAGGAGACTCA 25
||| ||||| ||||| ||||| |||||
Db 254 AAAGACACTTGCAGGAGATTCA 277

RESULT 37
US-09-012-692-122
; Sequence 122, Application US/09012692
; Patent No. 6214579
; GENERAL INFORMATION:
; APPLICANT: Grieve, Robert B.
; APPLICANT: Rushlow, Keith E.
; APPLICANT: Wu Hunter, Shirley
; APPLICANT: Frank, Glenn R.
; APPLICANT: Stiegler, Gary
; APPLICANT: Gaines, Patrick J.
; APPLICANT: Silver, Gary
; TITLE OF INVENTION: FLEA PROTEASE PROTEINS, NUCLEIC ACID
; TITLE OF INVENTION: MOLECULES AND USES THEREOF
; NUMBER OF SEQUENCES: 190
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheridan Ross & McIntosh
; STREET: 1700 Lincoln Street, Suite 3500
; CITY: Denver
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80203
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09012692
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/639,075
; FILING DATE: 24-APR-1996
; ATTORNEY/AGENT INFORMATION:

;
; APPLICATION NUMBER: US/09/012,692
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/639,075
; FILING DATE: 24-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Connell, Gary J.
; REGISTRATION NUMBER: 32,020
; REFERENCE/DOCKET NUMBER: 2618-25-C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 863-9700
; TELEFAX: (303) 863-0223
; INFORMATION FOR SEQ ID NO: 122:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 436 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2..436
; OTHER INFORMATION: /note= "At pos. bp 301, change A to
; OTHER INFORMATION: W; at pos. bp 342, change C to Y; at pos. bp 397, change C to
; OTHER INFORMATION: at pos. bp 431, change G to S. At pos. aa 100, 114 and 144,
; OTHER INFORMATION: substitute Xaa."
; US-09-012-692-122

Query Match 70.4%; Score 17.6; DB 3; Length 436;
Best Local Similarity 83.3%; Pred. No. 86;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AAAACACTTGAAGGAGACTCA 25
||| ||||| ||||| ||||| |||||
Db 254 AAAGACACTTGCAGGAGATTCA 277

RESULT 38
US-08-906-613-122
; Sequence 122, Application US/08906613
; Patent No. 6232096
; GENERAL INFORMATION:
; APPLICANT: Grieve, Robert B.
; APPLICANT: Rushlow, Keith E.
; APPLICANT: Wu Hunter, Shirley
; APPLICANT: Frank, Glenn R.
; APPLICANT: Stiegler, Gary
; APPLICANT: Gaines, Patrick J.
; APPLICANT: Silver, Gary
; TITLE OF INVENTION: FLEA PROTEASE PROTEINS, NUCLEIC ACID
; TITLE OF INVENTION: MOLECULES AND USES THEREOF
; NUMBER OF SEQUENCES: 190
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheridan Ross & McIntosh
; STREET: 1700 Lincoln Street, Suite 3500
; CITY: Denver
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80203
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/906,613
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/639,075
; FILING DATE: 24-APR-1996
; ATTORNEY/AGENT INFORMATION:

; NAME: Connell, Gary J.
; REGISTRATION NUMBER: 32,020
; REFERENCE/DOCKET NUMBER: 2618-25-C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 863-9700
; TELEFAX: (303) 863-0223
; INFORMATION FOR SEQ ID NO: 122:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 436 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2..436
; OTHER INFORMATION: /note= "at pos. bp 301, change A to
; OTHER INFORMATION: W; at pos. bp 342, change C to Y; at pos. bp 397, change C to
; OTHER INFORMATION: at pos. bp 431, change G to S. At pos. aa 100, 114 and 144,
; OTHER INFORMATION: substitute Xaa."
US-08-906-613-122

Query Match 70.4%; Score 17.6; DB 3; Length 436;
Best Local Similarity 83.3%; Pred. No. 86;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AAAAACTTGGAGGAGACTCA 25
||| ||||| ||||| ||||| |||||
Db 254 AAAGACACTTGCAGGAGATTCA 277

RESULT 39
US-09-949-016-168761/c
; Sequence 168761, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 168761
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-168761

Query Match 70.4%; Score 17.6; DB 3; Length 601;
Best Local Similarity 83.3%; Pred. No. 92;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAACTTGGAGGAGACTC 24
||| ||||| ||||| ||||| |||||
Db 47 AAAAAATAATTGTAATGGAGACTC 24

RESULT 40
US-09-513-999C-10285/c
; Sequence 10285, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.

; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59, US2, REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 10285
; LENGTH: 643
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 274
; OTHER INFORMATION: w=a or t
; US-09-513-999C-10285

Query Match 70.4%; Score 17.6; DB 3; Length 643;
Best Local Similarity 83.3%; Pred. No. 93;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAACTTGGAGGAGACTC 24
||| ||||| ||||| ||||| |||||
Db 632 AAAAACTTGGAGGAGTACCC 609

RESULT 41
US-08-881-094-13
; Sequence 13, Application US/08881094A
; Patent No. 6022739
; GENERAL INFORMATION:
; APPLICANT: Ryan, Clarence A
; APPLICANT: Pearce, Gregory L
; APPLICANT: McGurl, Barry F
; TITLE OF INVENTION: Systemin
; FILE REFERENCE: 7555-000001CPB
; CURRENT APPLICATION NUMBER: US/08/881,094A
; CURRENT FILING DATE: 1997-07-09
; EARLIER APPLICATION NUMBER: 08/308,887
; EARLIER FILING DATE: 1994-09-19
; EARLIER APPLICATION NUMBER: PCT/US93/02428
; EARLIER FILING DATE: 1993-03-18
; EARLIER APPLICATION NUMBER: 07/885,412
; EARLIER FILING DATE: 1992-03-19
; EARLIER APPLICATION NUMBER: 07/528,956
; EARLIER FILING DATE: 1990-05-25
; EARLIER APPLICATION NUMBER: PCT/US91/03685
; EARLIER FILING DATE: 1991-05-24
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 674
; TYPE: DNA
; ORGANISM: Solanum tuberosum
US-08-881-094-13

Query Match 70.4%; Score 17.6; DB 3; Length 674;
Best Local Similarity 83.3%; Pred. No. 94;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAACTTGGAGGAGACTC 24
||| ||||| ||||| ||||| |||||
Db 139 AAAAAATAATTGAAAGGAGACTC 162

RESULT 42
US-09-032-215-24
; Sequence 24, Application US/09032215
; Patent No. 6204010
; GENERAL INFORMATION:
; APPLICANT: Stiegler, Gary L.

```

RESULT 44
US-09-032-215-21
; Sequence 21, Application US/09032215
; Patent No. 6204010
; GENERAL INFORMATION:
; APPLICANT: Stiegler, Gary L.
; APPLICANT: Gaines, Patrick J.
; TITLE OF INVENTION: FLEA PROTEASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES, AND USES THEREOF
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheridan Ross P.C.
; STREET: 1700 Lincoln Street, Suite 3500
; CITY: Denver
; STATE: Colorado
; COUNTRY: U.S.A.
; ZIP: 80203
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

```

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII DOS TEXT
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/032,215
FILING DATE: 27-FEB-1998
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Connell, Gary J.
REGISTRATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 2618-25-C6
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 1303 nucleotides
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 149..1300
FEATURE:
NAME/KEY: W = A or T/U
LOCATION: 1165
FEATURE:
NAME/KEY: Y = C or T/U
LOCATION: 1206,1261
FEATURE:
NAME/KEY: S = C or G
LOCATION: 1295
FEATURE:
NAME/KEY: Xaa = Unknown
LOCATION: 339, 353, 383
US-09-032-215-21

Query Match 70.4%; Score 17.6; DB 3; Length 1303;
Best Local Similarity 83.3%; Pred. No. 1.1e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 2 AAAAAGACTTGGAGGGAGACTCA 25
DB 1118 AAAGACACTTGCAGGGAGATTCA 1141

RESULT 45
US-09-032-215-23/c
Sequence 23, Application US/09032215
Patent No. 6204010
GENERAL INFORMATION:
APPLICANT: Stiegler, Gary L.
APPLICANT: Gaines, Patrick J.
TITLE OF INVENTION: FLEA PROTEASE PROTEINS, NUCLEIC
TITLE OF INVENTION: ACID MOLECULES, AND USES THEREOF
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sheridan Ross P.C.
STREET: 1700 Lincoln Street, Suite 3500
CITY: Denver
STATE: Colorado
COUNTRY: U.S.A.
ZIP: 80203
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII DOS TEXT
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/032,215

FILING DATE: 27-FEB-1998
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Connell, Gary J.
REGISTRATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 2618-25-C6
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 1303 nucleotides
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: R = A or G
LOCATION: 43, 98
FEATURE:
NAME/KEY: W = A or T/U
LOCATION: 139
FEATURE:
NAME/KEY: S = C or G
LOCATION: 9
US-09-032-215-23

Query Match 70.4%; Score 17.6; DB 3; Length 1303;
Best Local Similarity 83.3%; Pred. No. 1.1e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 2 AAAAAGACTTGGAGGGAGACTCA 25
DB 186 AAAGACACTTGCAGGGAGATTCA 163

RESULT 46
US-09-023-655-188
Sequence 188, Application US/09023655
Patent No. 6607879
GENERAL INFORMATION:
APPLICANT: Cocks, Benjamin G.
APPLICANT: Susan G. Stuart
APPLICANT: Jeffrey J. Seilhamer
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
TITLE OF INVENTION: EXPRESSION
NUMBER OF SEQUENCES: 1508
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/023,655
FILING DATE: HEREWITH
CLASSIFICATION:
PRIOR APPLICATION NUMBER:
APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071

```
; REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 188:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2627 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: LUNGNOT01
; CLONE: 060309
US-09-023-655-188

Query Match 70.4%; Score 17.6; DB 3; Length 2627;
Best Local Similarity 83.3%; Pred. No. 1.2e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAACAACCTTGGAGGGAGACTC 24
Db 405 AAACAACACTTGGAAATTC 428

RESULT 47
US-09-620-312D-201
; Sequence 201, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yuning
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 201
; LENGTH: 4029
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (316)..(1842)
US-09-620-312D-201

Query Match 70.4%; Score 17.6; DB 3; Length 4029;
Best Local Similarity 83.3%; Pred. No. 1.3e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAACAACCTTGGAGGGAGACTC 24
Db 1803 AAACAACACTTGGAAATTC 1826

; REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 188:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2627 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: LUNGNOT01
; CLONE: 060309
US-09-023-655-188

Query Match 70.4%; Score 17.6; DB 3; Length 2627;
Best Local Similarity 83.3%; Pred. No. 1.2e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAACAACCTTGGAGGGAGACTC 24
Db 405 AAACAACACTTGGAAATTC 428

RESULT 48
US-09-949-016-15643/c
; Sequence 15643, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15643
; LENGTH: 13615
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-15643

Query Match 70.4%; Score 17.6; DB 3; Length 13615;
Best Local Similarity 83.3%; Pred. No. 1.7e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAACAACCTTGGAGGGAGACTC 24
Db 4021 AAAAACAACCTTGAAGAGTCACTC 3998

RESULT 49
US-09-949-016-12338/c
; Sequence 12338, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12338
; LENGTH: 26434
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-12338

Query Match 70.4%; Score 17.6; DB 3; Length 26434;
Best Local Similarity 83.3%; Pred. No. 1.9e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AAAAACAACCTTGGAGGGAGACTCA 25
Db 24563 AAAAACAACCTTGAAGAGATACCTTA 24540

RESULT 50
US-09-949-016-17324/c
; Sequence 17324, Application US/09949016
```

; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 17324
; LENGTH: 26434
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-17324

Query Match 70.4%; Score 17.6; DB 3; Length 26434;
Best Local Similarity 83.3%; Pred. No. 1.9e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AAAAAGCACTTGGAGGAGACTCA 25
|||||
Db 24563 AAAAGCACTTGGAGGAGACTTCA 24540

Search completed: February 3, 2006, 16:32:12
Job time : 87.6667 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 3, 2006, 14:37:28 ; Search time 365.556 Seconds
(without alignments)
565.535 Million cell updates/sec

Title: US-10-719-900-10

Perfect score: 25

Sequence: 1 aaaaacacttgaggagactca 25

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database : Published Applications NA Main:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	25	100.0	25	8	US-10-719-900-10
2	23.4	93.6	25	8	US-10-719-900-9
3	20.2	80.8	600	9	US-10-972-079-48833
C 4	19.8	79.2	1043	4	US-09-925-065A-91527
C 5	19.2	76.8	441	7	US-10-767-701-6654
C 6	19.2	76.8	873	7	US-10-152-319A-1905
7	18.8	75.2	1078	7	US-10-424-599-69851
C 8	18.6	74.4	837	4	US-09-925-065A-264666
C 9	18.6	74.4	893	4	US-09-925-065A-708512
10	18.6	74.4	1268	7	US-10-043-160-22
C 11	18.6	74.4	2511	3	US-09-823-245A-167
C 12	18.6	74.4	9210	3	US-09-764-877-3622
C 13	18.6	74.4	9210	3	US-09-764-877-3635
C 14	18.6	74.4	9210	6	US-10-242-515-3622
C 15	18.6	74.4	9210	6	US-10-242-515-3635
C 16	18.6	74.4	62169	7	US-10-323-696-157
C 17	18.4	73.6	114280	8	US-10-719-993-6808
C 18	18.2	72.8	188	6	US-10-029-386-23105
C 19	18.2	72.8	434	5	US-10-027-632-183294
C 20	18.2	72.8	434	6	US-10-027-632-183294
C 21	18.2	72.8	582	4	US-09-925-065A-548984
C 22	18.2	72.8	582	4	US-09-925-065A-548985
C 23	18.2	72.8	590	6	US-10-029-386-9405

24	18.2	72.8	786	3	US-09-938-842A-3478	Sequence 3478, Ap
25	18.2	72.8	786	3	US-09-938-842A-3478	Sequence 3478, Ap
C 26	18.2	72.8	865	5	US-10-198-846-7215	Sequence 7215, Ap
27	18.2	72.8	1260	6	US-10-369-493-33506	Sequence 33506, A
28	18.2	72.8	2092	3	US-09-968-007A-422	Sequence 422, App
29	18.2	72.8	2092	7	US-10-641-643-936	Sequence 936, App
30	18.2	72.8	2092	9	US-10-843-641A-6892	Sequence 6892, Ap
C 31	18.2	72.8	2306	6	US-10-369-493-46352	Sequence 46352, A
32	18.2	72.8	6721	7	US-10-693-999-19	Sequence 19, Appl
33	18.2	72.8	6787	9	US-10-887-553A-586	Sequence 586, App
34	17.8	71.2	368	8	US-10-425-115-116569	Sequence 116569, Sequence 423916,
35	17.8	71.2	490	4	US-09-925-065A-423916	Sequence 423917,
36	17.8	71.2	490	4	US-09-925-065A-423917	Sequence 4349, Ap
37	17.8	71.2	525	9	US-10-487-901-4349	Sequence 313637,
C 38	17.8	71.2	619	4	US-09-925-065A-313637	Sequence 313638,
C 39	17.8	71.2	619	4	US-09-925-065A-313638	Sequence 313639,
C 40	17.8	71.2	619	4	US-09-925-065A-313639	Sequence 724596,
41	17.8	71.2	631	4	US-09-925-065A-724596	Sequence 724597,
42	17.8	71.2	631	4	US-09-925-065A-724597	Sequence 724598,
43	17.8	71.2	631	4	US-09-925-065A-724598	Sequence 199, App
C 44	17.8	71.2	5347	5	US-10-205-823-199	Sequence 199, App
C 45	17.8	71.2	5347	10	US-11-051-454-199	Sequence 24, Appl
46	17.8	71.2	169998	6	US-10-380-931-24	Sequence 10, Appl
47	17.8	71.2	197496	3	US-09-877-177-10	Sequence 10, Appl
48	17.8	71.2	197496	6	US-10-426-836-10	Sequence 74, Appl
49	17.8	71.2	209083	9	US-10-461-862-74	Sequence 1, Appl
C 50	17.8	71.2	334462	9	US-10-496-011-1	Sequence 158180,
51	17.6	70.4	25	10	US-11-036-317-158180	Sequence 853249,
52	17.6	70.4	25	10	US-11-036-317-158180	Sequence 127517,
C 53	17.6	70.4	208	7	US-10-424-599-127517	Sequence 120024,
54	17.6	70.4	215	7	US-10-424-599-120024	Sequence 40695, A
55	17.6	70.4	287	7	US-10-424-599-40695	Sequence 6493, Ap
C 56	17.6	70.4	300	9	US-10-779-543-6493	Sequence 35405, A
57	17.6	70.4	311	7	US-10-242-535A-35405	Sequence 35405, A
58	17.6	70.4	311	7	US-10-085-783A-35405	Sequence 11595, A
C 59	17.6	70.4	389	8	US-10-779-543-11595	Sequence 7165, Ap
C 60	17.6	70.4	456	9	US-10-674-124A-7165	Sequence 30106, A
C 61	17.6	70.4	525	5	US-10-425-115-30106	Sequence 2547, Ap
C 62	17.6	70.4	525	5	US-10-066-543-2547	Sequence 171539,
C 63	17.6	70.4	561	4	US-09-925-065A-171539	Sequence 1995, Ap
C 64	17.6	70.4	563	3	US-09-969-034-1995	Sequence 1866, Ap
C 65	17.6	70.4	567	3	US-09-998-598-1866	Sequence 171540,
C 66	17.6	70.4	574	4	US-09-925-065A-171540	Sequence 187263,
C 67	17.6	70.4	585	5	US-10-027-632-187263	Sequence 187264,
C 68	17.6	70.4	585	5	US-10-027-632-187264	Sequence 187265,
C 69	17.6	70.4	585	5	US-10-027-632-187265	Sequence 187266,
C 70	17.6	70.4	585	5	US-10-027-632-187266	Sequence 408057,
C 71	17.6	70.4	585	6	US-10-027-632-187264	Sequence 677975,
C 72	17.6	70.4	585	6	US-10-027-632-187264	Sequence 1312, Ap
C 73	17.6	70.4	585	6	US-10-027-632-187265	Sequence 276, App
C 74	17.6	70.4	585	6	US-10-027-632-187266	Sequence 30110, A
C 75	17.6	70.4	586	4	US-09-925-065A-408057	Sequence 23368, A
76	17.6	70.4	786	7	US-10-152-319A-1512	Sequence 358, App
C 77	17.6	70.4	786	7	US-10-312-273-276	Sequence 1369, App
C 78	17.6	70.4	1110	6	US-10-425-115-30110	Sequence 744, App
C 79	17.6	70.4	1407	8	US-10-357-930-23368	Sequence 900, App
C 80	17.6	70.4	1607	5	US-10-116-802-358	Sequence 188, App
C 81	17.6	70.4	1966	6	US-10-062-674-1369	Sequence 116, App
C 82	17.6	70.4	2081	6	US-10-094-749-744	Sequence 4653, Ap
C 83	17.6	70.4	2477	6	US-10-094-749-744	Sequence 1018, Ap
C 84	17.6	70.4	2516	6	US-10-094-749-744	Sequence 201, App
85	17.6	70.4	2627	7	US-10-641-643-188	Sequence 201, App
C 86	17.6	70.4	3856	5	US-10-084-817-116	Sequence 5785, Ap
C 87	17.6	70.4	3878	9	US-10-956-157-4653	Sequence 34, Appl
88	17.6	70.4	4015	5	US-10-106-698-1018	Sequence 69, Appl
89	17.6	70.4	4029	5	US-10-037-270-201	Sequence 17574, A
90	17.6	70.4	4029	9	US-10-112-851-201	Sequence 17788, A
91	17.6	70.4	4029	9	US-10-112-851-201	
C 92	17.6	70.4	5184	3	US-09-764-891-5785	
93	17.6	70.4	94752	6	US-10-408-168-34	
94	17.6	70.4	140040	7	US-10-275-762-69	
95	17.6	70.4	156318	8	US-10-741-600-17574	
96	17.6	70.4	166472	8	US-10-741-600-17788	

```
97 17.6 70.4 174703 5 US-10-087-192-1336
98 17.6 70.4 247544 7 US-10-322-696-55
99 17.6 70.4 312477 7 US-10-317-888A-12
100 17.6 70.4 1230025 6 US-10-289-762-1
101 17.4 69.6 578 4 US-09-925-065A-330456
102 17.4 69.6 578 4 US-09-925-065A-330457
103 17.4 69.6 597 9 US-10-972-079-24527
104 17.4 69.6 600 9 US-10-972-079-24526
105 17.4 69.6 606 4 US-09-925-065A-330455
106 17.4 69.6 668 4 US-09-925-065A-647756
107 17.4 69.6 2682 3 US-09-938-842A-1842
108 17.4 69.6 2682 3 US-09-938-842A-1842
109 17.4 69.6 109559 7 US-10-322-281-137
110 17.2 68.8 300 9 US-10-779-543-3255
111 17.2 68.8 340 8 US-10-425-113-74382
112 17.2 68.8 402 9 US-10-779-543-11739
113 17.2 68.8 493 8 US-10-357-930-7162
114 17.2 68.8 525 4 US-09-925-065A-326219
115 17.2 68.8 589 4 US-09-925-065A-166948
116 17.2 68.8 596 4 US-09-925-065A-408056
117 17.2 68.8 599 9 US-10-972-079-84232
118 17.2 68.8 607 4 US-09-925-065A-727194
119 17.2 68.8 609 4 US-09-925-065A-646991
120 17.2 68.8 611 4 US-09-925-065A-154947
121 17.2 68.8 618 4 US-09-925-065A-836275
122 17.2 68.8 626 4 US-09-925-065A-554232
123 17.2 68.8 631 5 US-10-027-632-185105
124 17.2 68.8 631 5 US-10-027-632-185105
125 17.2 68.8 642 4 US-09-925-065A-780072
126 17.2 68.8 962 4 US-09-925-065A-705133
127 17.2 68.8 1082 4 US-09-925-065A-763
128 17.2 68.8 1082 4 US-09-925-065A-764
129 17.2 68.8 1082 4 US-09-925-065A-49681
130 17.2 68.8 1727 3 US-09-764-864-1611
131 17.2 68.8 2793 3 US-09-938-842A-394
132 17.2 68.8 2793 7 US-09-938-842A-394
133 17.2 68.8 2793 7 US-10-225-067-1
134 17.2 68.8 2793 7 US-10-374-780A-2561
135 17.2 68.8 3302 4 US-09-925-065A-553893
136 17.2 68.8 3526 7 US-10-437-963-60029
137 17.2 68.8 10368 10 US-11-097-143-5041
138 17.2 68.8 42863 6 US-10-017-161-1017
139 17.2 68.8 42863 6 US-10-292-798-865
140 17.2 68.8 325446 8 US-10-719-993-6824
141 17.2 68.8 374849 5 US-10-087-192-1627
142 17.2 68.8 383432 9 US-10-737-082-34
143 17.2 68.8 383432 9 US-10-765-790-34
144 17.2 68.8 1790242 8 US-10-719-993-6940
145 17 68.0 25 7 US-10-719-956-1799
146 17 68.0 25 7 US-10-719-956-1800
147 17 68.0 104 3 US-09-918-995-18212
148 17 68.0 112 3 US-09-918-995-18215
149 17 68.0 186 7 US-10-424-599-66385
150 17 68.0 201 7 US-10-741-601-870
```

ALIGNMENTS

```
RESULT 1
US-10-719-900-10
; Sequence 10, Application US/10719900
; Publication No. US20050026164A1
; GENERAL INFORMATION:
; APPLICANT: Xue Mei Zhou
; TITLE OF INVENTION: Methods of Genetic Analysis of Mouse
; FILE REFERENCE: 3528.1
; CURRENT APPLICATION NUMBER: US/10719,900
; PRIOR FILING DATE: 2003-11-20
; PRIOR APPLICATION NUMBER: 60/427,808
; NUMBER OF SEQ ID NOS: 2002 11 20
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1

; SEQ ID NO 10
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-719-900-10
Query Match 100.0%; Score 25; DB 8; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.14;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAAACACTTGGAGGGAGACTCA 25
|||||
Db 1 AAAAAACACTTGGAGGGAGACTCA 25

RESULT 2
US-10-719-900-9
; Sequence 9, Application US/10719900
; Publication No. US20050026164A1
; GENERAL INFORMATION:
; APPLICANT: Xue Mei Zhou
; TITLE OF INVENTION: Methods of Genetic Analysis of Mouse
; FILE REFERENCE: 3528.1
; CURRENT APPLICATION NUMBER: US/10719,900
; PRIOR FILING DATE: 2003-11-20
; PRIOR APPLICATION NUMBER: 60/427,808
; NUMBER OF SEQ ID NOS: 982914
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 9
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-719-900-9
Query Match 93.6%; Score 23.4; DB 8; Length 25;
Best Local Similarity 96.0%; Pred. No. 0.76;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAAAACACTTGGAGGGAGACTCA 25
|||||
Db 1 AAAAAACACTTGGAGGGAGACTCA 25

RESULT 3
US-10-972-079-48833
; Sequence 48833, Application US/10972079
; Publication No. US20050153317A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: ROSENFELD, David
; APPLICANT: KERR, Richard
; APPLICANT: BATES, Stephen
; APPLICANT: HOLM, Tom
; TITLE OF INVENTION: METHODS & SYSTEMS FOR INFERRING TRAITS TO BREED & MANAGE NON-BEEF
; FILE REFERENCE: MM1110-2
; CURRENT APPLICATION NUMBER: US/10/972,079
; CURRENT FILING DATE: 2004-10-22
; PRIOR APPLICATION NUMBER: US 60/514,333
; PRIOR FILING DATE: 2003-10-24
; NUMBER OF SEQ ID NOS: 96631
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 48833
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Chicken 19866894294418_1
US-10-972-079-48833

Query Match 80.8%; Score 20.2; DB 9; Length 600;
Best Local Similarity 88.0%; Pred. No. 35;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

QY 1 AAAAAACACTTGGGAAGGAGACTCA 25
|||||
Db 10 AAAAAACACTTGGGAAGGAGACTCA 34

RESULT 4
US-09-925-065A-91527/c
; Sequence 91527, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 91527
; LENGTH: 1043
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-91527

Query Match 79.2%; Score 19.8; DB 4; Length 1043;
Best Local Similarity 91.3%; Pred. No. 58;
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 AAAAAACACTTGGGAAGGAGACT 23
|||||
Db 367 AAAAAACACTTGGGAAGGAGACT 345

RESULT 5
US-10-767-701-6654/c
; Sequence 6654, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 6654
; LENGTH: 441
; TYPE: DNA
; ORGANISM: Sorghum bicolor
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(441)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS50873_1
US-10-767-701-6654

Query Match 76.8%; Score 19.2; DB 7; Length 441;
Best Local Similarity 87.5%; Pred. No. 98;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAAACACTTGGGAAGGAGACTCA 25
|||||
Db 363 AAAAAACACTTGGGAAGGAGACTAA 340

RESULT 6
US-10-152-319A-1905
; Sequence 1905, Application US/10152319A
; Publication No. US20040072160A1
; GENERAL INFORMATION:
; APPLICANT: Mendrick, Donna
; APPLICANT: Porter, Mark
; APPLICANT: Johnson, Kory
; APPLICANT: Higgs, Brandon
; APPLICANT: Castle, Arthur
; APPLICANT: Elashoff, Michael
; TITLE OF INVENTION: Molecular Toxicology Modeling
; FILE REFERENCE: 44921-5089-US
; CURRENT APPLICATION NUMBER: US/10/152,319A
; CURRENT FILING DATE: 2002-05-22
; PRIOR APPLICATION NUMBER: US 60/292,335
; PRIOR FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: US 60/297,523
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,925
; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: US 60/303,810
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/303,807
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/303,808
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/315,047
; PRIOR FILING DATE: 2001-08-28
; PRIOR APPLICATION NUMBER: US 60/324,928
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US 60/330,867
; PRIOR FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: US 60/330,462
; PRIOR FILING DATE: 2001-10-22
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2221
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1905
; LENGTH: 873
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. NM_031523
US-10-152-319A-1905

Query Match 76.8%; Score 19.2; DB 7; Length 873;
Best Local Similarity 87.5%; Pred. No. 1.1e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAAACACTTGGGAAGGAGACTCA 25
|||||
Db 658 AAAAAACACTTGGGAAGGAGACTCA 681

RESULT 7
US-10-424-599-69851/c
; Sequence 69851, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
US-10-424-599-69851/c

APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS
; FILE REFERENCE: GIN 6401
; CURRENT APPLICATION NUMBER: US/09/823,245A
; CURRENT FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/194,941
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 631
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 367
; LENGTH: 2511
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-823-245A-367

Query Match 74.4%; Score 18.6; DB 3; Length 2511;
Best Local Similarity 84.0%; Pred. No. 2.8e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAACACTTGGAGGAGACTCA 25
|||||
Db 1682 AAAACACATTGGAGGAGACTCA 1658

RESULT 12
US-09-764-877-3622/c
; Sequence 3622, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
; CURRENT APPLICATION NUMBER: US/09/764,877
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3622
; LENGTH: 9210
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-877-3622

Query Match 74.4%; Score 18.6; DB 3; Length 9210;
Best Local Similarity 84.0%; Pred. No. 2.8e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAACACTTGGAGGAGACTCA 25
|||||
Db 7719 AAAACACATTGGAGGAGACTCA 7695

RESULT 13
US-09-764-877-3635/c
; Sequence 3635, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
; CURRENT APPLICATION NUMBER: US/09/764,877
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3635
; LENGTH: 9210
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-877-3635

Query Match 74.4%; Score 18.6; DB 3; Length 9210;
Best Local Similarity 84.0%; Pred. No. 2.8e+02;

Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 AAAAAACACTTGGAGGAGACTCA 25
|||||
Db 7719 AAAACACATTGGAGGAGACTCA 7695

RESULT 14
US-10-242-515-3622/c
; Sequence 3622, Application US/10242515
; Publication No. US20040009488A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005C1
; CURRENT APPLICATION NUMBER: US/10/242,515
; CURRENT FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 09/764,877
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,758
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,963
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/217,496
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,447
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/218,290
; PRIOR FILING DATE: 2000-07-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3622
; LENGTH: 9210
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-242-515-3622

Query Match 74.4%; Score 18.6; DB 6; Length 9210;
Best Local Similarity 84.0%; Pred. No. 2.8e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAACACTTGGAGGAGACTCA 25
|||||
Db 7719 AAAACACATTGGAGGAGACTCA 7695

RESULT 15
US-10-242-515-3635/c
; Sequence 3635, Application US/10242515
; Publication No. US20040009488A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005C1
; CURRENT APPLICATION NUMBER: US/10/242,515
; CURRENT FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 09/764,877
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28

```
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,758
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,963
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/217,496
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,447
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/218,290
; PRIOR FILING DATE: 2000-07-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3635
; LENGTH: 9210
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-242-515-3635

Query Match          74.4%; Score 18.6; DB 6; Length 9210;
Best Local Similarity 84.0%; Pred. No. 2.8e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAACACTTGGAGGAGACTCA 25
    ||||| ||||| ||||| ||||| |||||
DB 7719 AAAAAACACTTGGAGGAGACTCA 7695

RESULT 16
US-10-322-696-157/c
; Sequence 157, Application US/10322696
; Publication No. US20040166490A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Malandro, Marc
; TITLE OF INVENTION: NOVEL THERAPEUTIC TARGETS IN CANCER
; FILE REFERENCE: 529452001200
; CURRENT APPLICATION NUMBER: US/10/322,696
; CURRENT FILING DATE: 2003-10-17
; NUMBER OF SEQ ID NOS: 186
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 157
; LENGTH: 62169
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(62169)
; OTHER INFORMATION: n = A,T,C or G
US-10-322-696-157

Query Match          74.4%; Score 18.6; DB 7; Length 62169;
Best Local Similarity 84.0%; Pred. No. 3.6e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAACACTTGGAGGAGACTCA 25
    ||||| ||||| ||||| ||||| |||||
DB 7740 AAAGAAACTTGGAGGGTCACTCA 7716

RESULT 17
US-10-719-993-6808
; Sequence 6808, Application US/10719993
; Publication No. US20040255849A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: ALZHEIMER'S DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001496
; CURRENT APPLICATION NUMBER: US/10/719,993
; CURRENT FILING DATE: 2003-11-24
```

```
; NUMBER OF SEQ ID NOS: 55342
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6808
; LENGTH: 114280
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-719-993-6808

Query Match          73.6%; Score 18.4; DB 8; Length 114280;
Best Local Similarity 95.0%; Pred. No. 4.8e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 AAAACACTTGGAGGAGACT 23
    ||||| ||||| ||||| ||||| |||||
DB 82496 AAACAATTGGAGGAGACT 82515

RESULT 18
US-10-029-386-23105/c
; Sequence 23105, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 23105
; LENGTH: 188
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR2.3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.89
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.98
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.2
; OTHER INFORMATION: NT HIT: gi4503164, EVALUE 1.00e-79
; OTHER INFORMATION: SWISSPROT HIT: Q9JLV5, EVALUE 3.00e-26
; OTHER INFORMATION: EST_HUMAN HIT: BF038331.1, EVALUE 2.00e-79
US-10-029-386-23105

Query Match          72.8%; Score 18.2; DB 6; Length 188;
Best Local Similarity 87.0%; Pred. No. 2.5e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAACACTTGGAGGAGACT 23
    ||||| ||||| ||||| ||||| |||||
DB 86 AAACAACACTTGGCAAGGAGACT 64

RESULT 19
US-10-027-632-183294/c
; Sequence 183294, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
```

```

, PRIOR APPLICATION NUMBER: US 60/185,218
, PRIOR FILING DATE: 2000-02-24
, PRIOR APPLICATION NUMBER: US 60/167,363
, PRIOR FILING DATE: 1999-11-23
, PRIOR APPLICATION NUMBER: US 60/156,358
, PRIOR FILING DATE: 1999-09-28
, PRIOR APPLICATION NUMBER: US 60/146,002
, PRIOR FILING DATE: 1999-08-09
, NUMBER OF SEQ ID NOS: 325720
, SOFTWARE: FastSeq for Windows Version 4.0
, SEQ ID NO 183294
, LENGTH: 434
, TYPE: DNA
, ORGANISM: Human
US-10-027-632-183294

```

Query Match 72.8%; Score 18.2; DB 5; Length 434;
Best Local Similarity 87.0%; Pred. No. 2.8e+02;
Matches 20; Conservative 0; Mismatches 3; Indels

QY 1 AAAAAACACTTGAAGCGAGCT 23
||||| ||||| ||||| ||||| ||
Db 111 AAAAAATACCTTGAAGCGAGCT 89

RESULTS

US-10-027-632-183294/c
; Sequence 183294, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:

APPLICANT: Wang, David G.

; TITLE OF INVENTION: Identification and Mapping of Sing
 ; TITLE OF INVENTION: Polymorphisms in the Human Genome
 ; TITLE OF INVENTION: Polymorphisms in the Human Genome

```

FILE OF INVENTION: Polymorphisms in the
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US 10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ IDS: 325720
SOFTWARE: FASTSEQ for Windows Version 4.0.C
SEQ ID NO 183294
LENGTH: 434
TYPE: DNA
ORGANISM: Human
US-10-027-632-183294

```

Query Match	72.8%	Score 18.2;	DB 6;	Length 434;
Best Local Similarity	87.0%;	Pred. No. 2.8s+02;		
Matches	20. Conservative	0. Mismatches	3. Indels	0. Gaps

Qy 1 AAAAAACACTTGAAGGGAGACT 23
|||
Db 111 AAAAAATACCTGTAAAGGGAGCT 89

RESULT 21

RESULTS 21
US-09-925-065A-548984
; Sequence 548984, Application US/09325065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.

```

; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108927.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 548984
; LENGTH: 582
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-925-065A-548984

```

Query Match	72.8%	Score 18.2	DB 4	Length 582
-------------	-------	------------	------	------------

Query Match	Best Local Similarity	87.0%;	Pred. No. 2.9e+02;	0. Mismatches	3. Indels	0. Gaps	0.
Query Match	Score 10.2;	DB 4;	12.8%;	Score 10.2;	DB 4;	12.8%;	Score 10.2;

Qy 1 AAAAAACACTTCGAAGGGAGACT 23
|||||
Db 417 AAAAAACACTTCAGGGGGAGACT 439
|||||

RESULT 22

US-09-925-065A-548985
; Sequence 548985, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:

```

1  APPLICANT: Wang, David G.
2  TITLE OF INVENTION: Identification and Mapping of Single
3  NUCLEOTIDE POLYMORPHISMS IN THE HUMAN GENOME
4  FILE REFERENCE: 108827.135
5  CURRENT APPLICATION NUMBER: US/09/925,065A
6  CURRENT FILING DATE: 2001-08-08
7  PRIOR APPLICATION NUMBER: US 60/243,096
8  PRIOR FILING DATE: 2000-10-24
9  PRIOR APPLICATION NUMBER: US 60/252,147
10 PRIOR FILING DATE: 2000-11-20
11 PRIOR APPLICATION NUMBER: US 60/250,092
12 PRIOR FILING DATE: 2000-11-30
13 PRIOR APPLICATION NUMBER: US 60/261,766
14 PRIOR FILING DATE: 2001-01-16
15 PRIOR APPLICATION NUMBER: US 60/289,846
16 PRIOR FILING DATE: 2001-05-09
17 NUMBER OF SEQ ID NOS: 957086
18 SOFTWARE: FastSEQ for Windows Version 4.0
19 SEQ ID NO 548985
20 LENGTH: 582

```

ORGANISM: Homo sapiens
UHS-09-925-065A-548985

Query Match	72.8%;	Score 18.2;	DB 4;	Length 582;
Best Local Similarity	87.0%;	Pred. No. 2.9e+02;		
Matches	20;	Mismatches	3;	Indels
	Conservative			

QY 1 AAAAAACACTTGGAAAGGAGACT 23
|||||
pb 417 AAAAAACACTTTCAGGGAGACT 439
|||||

RESULT 23

RESULT 23
US-10-029-386-9405/c
: Sequence 9405: Application US/100293386

US-09-938-842A-3478

Sequence 3478, Application US/09938842A
Publication No. US20040009476A9
GENERAL INFORMATION:
APPLICANT: Harper, Jeff
APPLICANT: Kreps, Joel
APPLICANT: Wang, Xun
APPLICANT: Zhu, Tong
TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
TITLE OF INVENTION: SAME, AND METHODS OF USE
FILE REFERENCE: SCRIPI300-3
CURRENT APPLICATION NUMBER: US/09/938,842A
CURRENT FILING DATE: 2001-08-24
PRIOR APPLICATION NUMBER: US 60/227,866
PRIOR FILING DATE: 2000-08-24
PRIOR APPLICATION NUMBER: US 60/264,647
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/300,111
PRIOR FILING DATE: 2001-06-22
NUMBER OF SEQ ID NOS: 5379
SEQ ID NO 3478
LENGTH: 786
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-09-938-842A-3478

Query Match 72.8%; Score 18.2; DB 3; Length 786;
Best Local Similarity 87.0%; Pred. No. 3.1e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAAACACTTGGAGGGAGACTC 24
|||||

DB 512 ATAAACACTTGGAGAGAGATC 534
|||||

RESULT 26

US-10-198-846-7215/c

Sequence 7215, Application US/10198846
Publication No. US2003009974A1
GENERAL INFORMATION:
APPLICANT: Xu, Yongyao
APPLICANT: Wang, Youzhen
APPLICANT: Steinmann, Kathleen
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
TITLE OF INVENTION: THERAPY OF BREAST CANCER
FILE REFERENCE: MRI-049
CURRENT APPLICATION NUMBER: US/10/198,846
CURRENT FILING DATE: 2002-07-18
PRIOR APPLICATION NUMBER: 60/306,220
PRIOR FILING DATE: 2001-07-18
NUMBER OF SEQ ID NOS: 14084
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 7215
LENGTH: 865
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc.feature
LOCATION: 78, 216, 238, 246, 250, 252, 283, 299, 439, 451, 481, 525,
LOCATION: 532, 614, 622, 625, 645, 646, 663, 675, 689, 713, 718, 723,
LOCATION: 729, 739, 740, 756, 766, 771, 794, 810, 816, 820, 829, 836,
LOCATION: 838, 849, 856, 859, 862
OTHER INFORMATION: n = A,T,C or G
US-10-198-846-7215

Query Match 72.8%; Score 18.2; DB 5; Length 865;
Best Local Similarity 87.0%; Pred. No. 3.1e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAACACTTGGAGGGAGACT 23
|||||

US-09-938-842A-3478

Sequence 3478, Application US/09938842A
Patent No. US20020160378A1
GENERAL INFORMATION:
APPLICANT: Harper, Jeff
APPLICANT: Kreps, Joel
APPLICANT: Wang, Xun
APPLICANT: Zhu, Tong
TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
TITLE OF INVENTION: SAME, AND METHODS OF USE
FILE REFERENCE: SCRIPI300-3
CURRENT APPLICATION NUMBER: US/09/938,842A
CURRENT FILING DATE: 2001-08-24
PRIOR APPLICATION NUMBER: US 60/227,866
PRIOR FILING DATE: 2000-08-24
PRIOR APPLICATION NUMBER: US 60/264,647
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/300,111
PRIOR FILING DATE: 2001-06-22
NUMBER OF SEQ ID NOS: 5379
SEQ ID NO 3478
LENGTH: 786
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-09-938-842A-3478

Query Match 72.8%; Score 18.2; DB 3; Length 786;
Best Local Similarity 87.0%; Pred. No. 3.1e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAAACACTTGGAGGGAGACTC 24
|||||

DB 512 ATAAACACTTGGAGAGAGATC 534
|||||

RESULT 25

US-09-938-842A-3478

Sequence 3478, Application US/09938842A
Patent No. US20020160378A1
GENERAL INFORMATION:
APPLICANT: Harper, Jeff
APPLICANT: Kreps, Joel
APPLICANT: Wang, Xun
APPLICANT: Zhu, Tong
TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
TITLE OF INVENTION: SAME, AND METHODS OF USE
FILE REFERENCE: SCRIPI300-3
CURRENT APPLICATION NUMBER: US/09/938,842A
CURRENT FILING DATE: 2001-08-24
PRIOR APPLICATION NUMBER: US 60/227,866
PRIOR FILING DATE: 2000-08-24
PRIOR APPLICATION NUMBER: US 60/264,647
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/300,111
PRIOR FILING DATE: 2001-06-22
NUMBER OF SEQ ID NOS: 5379
SEQ ID NO 3478
LENGTH: 786
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-09-938-842A-3478

Query Match 72.8%; Score 18.2; DB 3; Length 786;
Best Local Similarity 87.0%; Pred. No. 3.1e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAAACACTTGGAGGGAGACTC 24
|||||

DB 512 ATAAACACTTGGAGAGAGATC 534
|||||

RESULT 25

US-09-938-842A-3478

Sequence 3478, Application US/09938842A
Patent No. US20020160378A1
GENERAL INFORMATION:
APPLICANT: Harper, Jeff
APPLICANT: Kreps, Joel
APPLICANT: Wang, Xun
APPLICANT: Zhu, Tong
TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
TITLE OF INVENTION: SAME, AND METHODS OF USE
FILE REFERENCE: SCRIPI300-3
CURRENT APPLICATION NUMBER: US/09/938,842A
CURRENT FILING DATE: 2001-08-24
PRIOR APPLICATION NUMBER: US 60/227,866
PRIOR FILING DATE: 2000-08-24
PRIOR APPLICATION NUMBER: US 60/264,647
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/300,111
PRIOR FILING DATE: 2001-06-22
NUMBER OF SEQ ID NOS: 5379
SEQ ID NO 3478
LENGTH: 786
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-09-938-842A-3478

Query Match 72.8%; Score 18.2; DB 3; Length 786;
Best Local Similarity 87.0%; Pred. No. 3.1e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAAACACTTGGAGGGAGACTC 24
|||||

DB 512 ATAAACACTTGGAGAGAGATC 534
|||||

RESULT 25

US-09-938-842A-3478

Sequence 3478, Application US/09938842A
Patent No. US20020160378A1
GENERAL INFORMATION:
APPLICANT: Harper, Jeff
APPLICANT: Kreps, Joel
APPLICANT: Wang, Xun
APPLICANT: Zhu, Tong
TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
TITLE OF INVENTION: SAME, AND METHODS OF USE
FILE REFERENCE: SCRIPI300-3
CURRENT APPLICATION NUMBER: US/09/938,842A
CURRENT FILING DATE: 2001-08-24
PRIOR APPLICATION NUMBER: US 60/227,866
PRIOR FILING DATE: 2000-08-24
PRIOR APPLICATION NUMBER: US 60/264,647
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/300,111
PRIOR FILING DATE: 2001-06-22
NUMBER OF SEQ ID NOS: 5379
SEQ ID NO 3478
LENGTH: 786
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-09-938-842A-3478

Query Match 72.8%; Score 18.2; DB 3; Length 786;
Best Local Similarity 87.0%; Pred. No. 3.1e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAAACACTTGGAGGGAGACTC 24
|||||

DB 512 ATAAACACTTGGAGAGAGATC 534
|||||

RESULT 25

US-09-938-842A-3478

Sequence 3478, Application US/09938842A
Patent No. US20020160378A1
GENERAL INFORMATION:
APPLICANT: Harper, Jeff
APPLICANT: Kreps, Joel
APPLICANT: Wang, Xun
APPLICANT: Zhu, Tong
TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
TITLE OF INVENTION: SAME, AND METHODS OF USE
FILE REFERENCE: SCRIPI300-3
CURRENT APPLICATION NUMBER: US/09/938,842A
CURRENT FILING DATE: 2001-08-24
PRIOR APPLICATION NUMBER: US 60/227,866
PRIOR FILING DATE: 2000-08-24
PRIOR APPLICATION NUMBER: US 60/264,647
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/300,111
PRIOR FILING DATE: 2001-06-22
NUMBER OF SEQ ID NOS: 5379
SEQ ID NO 3478
LENGTH: 786
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-09-938-842A-3478

Query Match 72.8%; Score 18.2; DB 3; Length 786;
Best Local Similarity 87.0%; Pred. No. 3.1e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAAACACTTGGAGGGAGACTC 24
|||||

DB 512 ATAAACACTT

Db 575 AAAAAAGACTTGGAGGTAAC 553

RESULT 27

US-10-369-493-33506
; Sequence 33506, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 33506
; LENGTH: 1260
; TYPE: DNA
; ORGANISM: Desulfitobacterium hafniense
US-10-369-493-33506

Query Match 72.8%; Score 18.2; DB 6; Length 1260;
Best Local Similarity 87.0%; Pred. No. 3.3e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AAAAAACACTTGGAGGGAGACT 23
Db 25 AAAAAACAGTTGGAAGGTAGACT 47

RESULT 28

US-09-968-007A-422
; Sequence 422, Application US/09968007A
; Publication No. US20040115625A1
; GENERAL INFORMATION:
; APPLICANT: Ebner, Reinhard
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signa
; FILE REFERENCE: 689290-71
; CURRENT APPLICATION NUMBER: US/09/968,007A
; CURRENT FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US/60/237,172
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: US/60/237,173
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: US/60/237,278
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: US/60/237,294
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: US/60/237,295
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: US/60/237,316
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 1001
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 422
; LENGTH: 2092
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-968-007A-422

Query Match 72.8%; Score 18.2; DB 3; Length 2092;
Best Local Similarity 87.0%; Pred. No. 3.5e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AAAAAACACTTGGAGGGAGACT 23

Db 724 AAAAAACACTTGGCAAGGAGACT 746

RESULT 29

US-10-641-643-936
; Sequence 936, Application US/10641643
; Publication No. US20040077003A1
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESSER: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/641,643
; FILING DATE: 14-Aug-2003
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: <Unknown>
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 936:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2092 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: G1381145
; SEQUENCE DESCRIPTION: SEQ ID NO: 936 :

US-10-641-643-936

Query Match 72.8%; Score 18.2; DB 7; Length 2092;
Best Local Similarity 87.0%; Pred. No. 3.5e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AAAAAACACTTGGAGGGAGACT 23

Db 724 AAAAAACACTTGGCAAGGAGACT 746

RESULT 30

US-10-843-641A-6892
; Sequence 6892, Application US/10843641A
; Publication No. US20050064454A1
; GENERAL INFORMATION:
; APPLICANT: Avalon Pharmaceuticals, Inc.
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
; TITLE OF INVENTION: Signature Gene Sets
; FILE REFERENCE: 689290-189
; CURRENT APPLICATION NUMBER: US/10/843,641A
; CURRENT FILING DATE: 2004-05-12
; PRIOR APPLICATION NUMBER: US/09/873,367

```
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US/09/954,531
; PRIOR FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/09/954,456
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/962,436
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/962,832
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/964,824
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US/09/967,768
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US/09/968,007
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US/09/969,347
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US/09/969,708
; PRIOR FILING DATE: 2001-10-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 8447
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6892
; LENGTH: 2092
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(2092)
; OTHER INFORMATION: n=a,t,g or c
US-10-843-641A-6892

Query Match 72.8%; Score 18.2; DB 9; Length 2092;
Best Local Similarity 87.0%; Pred. No. 3.5e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAACACTTGGAGGGGAGACT 23
DB 724 AAACACACTTGGCAAGGAGACT 746

RESULT 31
US-10-369-493-46352/c
; Sequence 46352, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 46352
; LENGTH: 2306
; TYPE: DNA
; ORGANISM: Schizosaccharomyces pombe
US-10-369-493-46352

Query Match 72.8%; Score 18.2; DB 6; Length 2306;
Best Local Similarity 87.0%; Pred. No. 3.5e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAACACTTGGAGGGGAGACT 24
DB 1161 ACAACACTAGGAGGAGGCTC 1139
```

```
RESULT 32
US-10-693-999-19
; Sequence 19, Application US/10693999
; Publication No. US20040137597A1
; GENERAL INFORMATION:
; APPLICANT: Meso Scale Technologies, LLC
; APPLICANT: Davydov, Ilia
; APPLICANT: Kenten, John H.
; APPLICANT: Biebuyck, Hans
; APPLICANT: Oberoi, Pankaj
; TITLE OF INVENTION: THE UBIQUITYLATION AND METHODS FOR MEASURING
; TITLE OF INVENTION: SUBSTRATES OF N-END RULE UBIQUITYLATION OF THESE SUBSTRATES
; FILE REFERENCE: 2528-10 / P14050US0
; CURRENT APPLICATION NUMBER: US/10/693,999
; CURRENT FILING DATE: 2003-10-28
; PRIOR APPLICATION NUMBER: US 60/422,448
; PRIOR FILING DATE: 2002-10-30
; PRIOR APPLICATION NUMBER: US 60/486,529
; PRIOR FILING DATE: 2003-07-19
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 19
; LENGTH: 6721
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-693-999-19

Query Match 72.8%; Score 18.2; DB 7; Length 6721;
Best Local Similarity 87.0%; Pred. No. 4.1e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAACACTTGGAGGGGAGACT 23
DB 1618 AAACACACTTGGCAAGGAGACT 1640

RESULT 33
US-10-887-553A-586
; Sequence 586, Application US/10887553A
; Publication No. US20050085436A1
; GENERAL INFORMATION:
; APPLICANT: Garza, Dan
; APPLICANT: Li, Hao
; TITLE OF INVENTION: Method to treat conditions associated
; TITLE OF INVENTION: with insulin signalling dysregulation
; FILE REFERENCE: 4-33262
; CURRENT APPLICATION NUMBER: US/10/887,553A
; CURRENT FILING DATE: 2004-07-08
; PRIOR APPLICATION NUMBER: 60/485,883
; PRIOR FILING DATE: 2003-08-07
; NUMBER OF SEQ ID NOS: 1208
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 586
; LENGTH: 6787
; TYPE: DNA
; ORGANISM: human
US-10-887-553A-586

Query Match 72.8%; Score 18.2; DB 9; Length 6787;
Best Local Similarity 87.0%; Pred. No. 4.1e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAACACTTGGAGGGGAGACT 23
DB 1685 AAACACACTTGGCAAGGAGACT 1707

RESULT 34
US-10-425-115-116569
; Sequence 116569, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
```

```
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 116569
; LENGTH: 368
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_3779C.1
; US-10-425-115-116569

Query Match          71.2%; Score 17.8; DB 8; Length 368;
Best Local Similarity 90.5%; Pred. No. 4.2e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAACACTTGGAGGGAGA 21
   ||||| ||||| ||||| |||||
Db 168 AAAAAACACTTAGGAAGGGGA 188

RESULT 35
US-09-925-065A-423916
; Sequence 423916, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 423916
; LENGTH: 490
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-925-065A-423916

Query Match          71.2%; Score 17.8; DB 4; Length 490;
Best Local Similarity 90.5%; Pred. No. 4.4e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAACACTTGGAGGGAGA 21
   ||||| ||||| ||||| |||||
Db 355 AAAAAACACTTGGAGGGAGA 375

RESULT 36
US-09-925-065A-423917
; Sequence 423917, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 423916
; LENGTH: 490
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-925-065A-423916

Query Match          71.2%; Score 17.8; DB 4; Length 490;
Best Local Similarity 90.5%; Pred. No. 4.4e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAACACTTGGAGGGAGA 21
   ||||| ||||| ||||| |||||
Db 355 AAAAAACACTTGGAGGGAGA 375

RESULT 37
US-10-487-901-4349
; Sequence 4349, Application US/10487901
; Publication No. US20050091708A1
; GENERAL INFORMATION:
; APPLICANT: Oreido, Jeremiah Vincent
; APPLICANT: McCreery, David
; APPLICANT: Pell, Randy
; APPLICANT: Miller, Barbara
; APPLICANT: Weglarz, Thaddeus
; APPLICANT: Gachotte, Daniel
; APPLICANT: Blakeslee, Beth
; APPLICANT: Larrinua, Ignacio
; APPLICANT: Reddy, Avutu
; APPLICANT: Shukla, Vipula
; APPLICANT: Crosley, Rodney
; TITLE OF INVENTION: Nucleic Acid Compositions Conferring Altered Metabolic Character
; FILE REFERENCE: DOW-08552
; CURRENT APPLICATION NUMBER: US/10/487,901
; CURRENT FILING DATE: 2004-02-26
; NUMBER OF SEQ ID NOS: 7560
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 4349
; LENGTH: 525
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic
; US-10-487-901-4349

Query Match          71.2%; Score 17.8; DB 9; Length 525;
Best Local Similarity 90.5%; Pred. No. 4.4e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 AACACTTGGAGGGAGACTCA 25
   ||||| ||||| ||||| |||||
Db 284 AACACTTGGATGAGAGACTCA 304

RESULT 38
US-09-925-065A-313637/c
; Sequence 313637, Application US/09925065A
; Publication No. US20050228172A9
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
```

```
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US 09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 313637
; LENGTH: 619
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-313637
```

```
Query Match 71.2%; Score 17.8; DB 4; Length 619;
Best Local Similarity 90.5%; Pred. No. 4.5e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 1 AAAAAACACTTGGGAAGGAGA 21
||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 297 AAAAAACACTTGGGAAGGAGA 277
```

RESULT 39

```
US-09-925-065A-313638/c
; Sequence 313638, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US 09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 313638
; LENGTH: 619
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-313638
```

```
Query Match 71.2%; Score 17.8; DB 4; Length 619;
Best Local Similarity 90.5%; Pred. No. 4.5e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 1 AAAAAACACTTGGGAAGGAGA 21
||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 297 AAAAAACACTTGGGAAGGAGA 277
```

RESULT 40

```
US-09-925-065A-313639/c
; Sequence 313639, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US 09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 313639
; LENGTH: 619
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-313639
```

```
Query Match 71.2%; Score 17.8; DB 4; Length 619;
Best Local Similarity 90.5%; Pred. No. 4.5e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 1 AAAAAACACTTGGGAAGGAGA 21
||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 297 AAAAAACACTTGGGAAGGAGA 277
```

RESULT 41

```
US-09-925-065A-724596
; Sequence 724596, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US 09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 724596
; LENGTH: 631
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-724596
```

```
Query Match 71.2%; Score 17.8; DB 4; Length 631;
Best Local Similarity 90.5%; Pred. No. 4.5e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 1 AAAAAACACTTGGGAAGGAGA 21
||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 284 AAAAAACACTTGGGAATGGAAA 304
```

RESULT 42
US-09-925-065A-724597
; Sequence 724597, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT FILING DATE: 2001-08-08
; CURRENT APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 724597
; LENGTH: 631
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-724597

Query Match 71.2%; Score 17.8; DB 4; Length 631;
Best Local Similarity 90.5%; Pred. No. 4.5e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAACACTTGGAGGAGA 21
|||||

Db 284 AAAAAACACTTGGATGGAAA 304

RESULT 43
US-09-925-065A-724598
; Sequence 724598, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 724598
; LENGTH: 631
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-724598

Query Match 71.2%; Score 17.8; DB 4; Length 631;
Best Local Similarity 90.5%; Pred. No. 4.5e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAACACTTGGAGGAGA 21
|||||

Db 284 AAAAAACACTTGGATGGAAA 304

RESULT 44
US-10-205-823-199/c
; Sequence 199, Application US/10205823
; Publication No. US20030108963A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Endege, Wilson O.
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Gorbacheva, Bella
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Monsey, Angela M.
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Anderson, Dustin
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; THERAPY OF PROSTATE CANCER
; FILE REFERENCE: MRI-044
; CURRENT APPLICATION NUMBER: US/10/205,823
; CURRENT FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: 60/307,982
; PRIOR FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: 60/314,356
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/325,020
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: 60/341,746
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: 60/362,158
; PRIOR FILING DATE: 2002-03-05
; NUMBER OF SEQ ID NOS: 455
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 199
; LENGTH: 5347
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-205-823-199

Query Match 71.2%; Score 17.8; DB 5; Length 5347;
Best Local Similarity 90.5%; Pred. No. 6.1e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAACACTTGGAGGAGA 21
|||||

Db 1038 AAAAAACACTTGGAGGAGA 1018

RESULT 45
US-11-051-454-199/c
; Sequence 199, Application US/11051454
; Publication No. US20050191673A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Endege, Wilson O.
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Gorbacheva, Bella
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Monsey, Angela M.
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Anderson, Dustin
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; THERAPY OF PROSTATE CANCER
; FILE REFERENCE: MRI-044
; CURRENT APPLICATION NUMBER: US/11/051,454
; CURRENT FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: 60/307,982
; PRIOR FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: 60/314,356
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/325,020
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: 60/341,746
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: 60/362,158
; PRIOR FILING DATE: 2002-03-05
; NUMBER OF SEQ ID NOS: 455
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 199
; LENGTH: 5347
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-051-454-199/c

Query Match 71.2%; Score 17.8; DB 5; Length 5347;
Best Local Similarity 90.5%; Pred. No. 6.1e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
; FILE REFERENCE: MRI-044
; CURRENT APPLICATION NUMBER: US/11/051,454
; CURRENT FILING DATE: 2005-02-04
; PRIOR APPLICATION NUMBER: US/10/205,823
; PRIOR FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: 60/307,982
; PRIOR FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: 60/314,356
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/325,020
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: 60/341,746
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: 60/362,158
; PRIOR FILING DATE: 2002-03-05
; NUMBER OF SEQ ID NOS: 455
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 199
; LENGTH: 5347
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-051-454-199

Query Match          71.2%; Score 17.8; DB 10; Length 5347;
Best Local Similarity 90.5%; Pred. No. 6.1e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAACACTTGGAGGGGAGA 21
   ||||| ||||| |||||
Db 1038 AAGAAACACATGGAAGGGAGA 1018

RESULT 46
US-10-380-931-24
; Sequence 24, Application US/10380931
; Publication No. US20030215944A1
; GENERAL INFORMATION:
; APPLICANT: Isis Pharmaceuticals, Inc.
; APPLICANT: C. Frank Bennett
; APPLICANT: Jacqueline Wyatt
; APPLICANT: Susan M. Freier
; TITLE OF INVENTION: OLIGONUCLEOTIDE INHIBITION OF HER-1 EXPRESSION
; FILE REFERENCE: RTSP-0187
; CURRENT APPLICATION NUMBER: US/10/380,931
; CURRENT FILING DATE: 2003-03-18
; PRIOR APPLICATION NUMBER: 09/676,610
; PRIOR FILING DATE: 2000-09-29
; NUMBER OF SEQ ID NOS: 182
; SEQ ID NO 24
; LENGTH: 169998
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: exon
; LOCATION: (1208)...(1472)
; FEATURE:
; NAME/KEY: intron
; LOCATION: (1473)...(124390)
; FEATURE:
; NAME/KEY: exon
; LOCATION: (124391)...(124544)
; FEATURE:
; NAME/KEY: intron
; LOCATION: (124545)...(125409)
; FEATURE:
; NAME/KEY: exon
; LOCATION: (125410)...(125595)
; FEATURE:
; NAME/KEY: intron
; LOCATION: (125596)...(128711)
; FEATURE:
; NAME/KEY: exon
; LOCATION: (128712)...(128848)
; FEATURE:
; NAME/KEY: intron
; LOCATION: (128849)...(133400)
; FEATURE:
; NAME/KEY: exon
; LOCATION: (133401)...(133469)
; FEATURE:
; NAME/KEY: intron
; LOCATION: (133470)...(134652)
; FEATURE:
; NAME/KEY: exon
; LOCATION: (134653)...(134773)
; FEATURE:
; NAME/KEY: intron
; LOCATION: (134774)...(136116)
; FEATURE:
; NAME/KEY: exon
; LOCATION: (136117)...(136261)
; FEATURE:
; NAME/KEY: intron
; LOCATION: (136262)...(137936)
; FEATURE:
; NAME/KEY: exon
; LOCATION: (137937)...(138053)
; FEATURE:
; NAME/KEY: intron
; LOCATION: (138054)...(138637)
; FEATURE:
; NAME/KEY: exon
; LOCATION: (138638)...(138766)
; FEATURE:
; NAME/KEY: intron
; LOCATION: (138767)...(138864)
; FEATURE:
; NAME/KEY: exon
; LOCATION: (138865)...(138940)
; FEATURE:
; NAME/KEY: intron
; LOCATION: (138941)...(139765)
; FEATURE:
; NAME/KEY: exon
; LOCATION: (139766)...(139860)
; FEATURE:
; NAME/KEY: intron
; LOCATION: (139861)...(142245)
; FEATURE:
; NAME/KEY: exon
; LOCATION: (142246)...(142445)
; FEATURE:
; NAME/KEY: intron
; LOCATION: (142446)...(143605)
; FEATURE:
; NAME/KEY: exon
; LOCATION: (143606)...(143738)
; FEATURE:
; NAME/KEY: intron
; LOCATION: (143739)...(145838)
; FEATURE:
; NAME/KEY: exon
; LOCATION: (145839)...(145931)
; FEATURE:
; NAME/KEY: intron
; LOCATION: (145932)...(147385)
; FEATURE:
; NAME/KEY: exon
; LOCATION: (147386)...(147544)
; FEATURE:
; NAME/KEY: intron
; LOCATION: (147545)...(153274)
; FEATURE:
; NAME/KEY: exon
; LOCATION: (153275)...(153321)
; FEATURE:
```

```

; NAME/KEY: intron
; LOCATION: (153322)...(155088)
; FEATURE:
; NAME/KEY: exon
; LOCATION: (155089)...(155231)
; FEATURE:
; NAME/KEY: intron
; LOCATION: (155232)...(156025)
; FEATURE:
; NAME/KEY: exon
; LOCATION: (156026)...(156151)
; FEATURE:
; NAME/KEY: intron
; LOCATION: (156152)...(156826)
; FEATURE:
; NAME/KEY: exon
; LOCATION: (156827)...(156928)
; FEATURE:
; NAME/KEY: intron
; LOCATION: (156929)...(163399)
; FEATURE:
; NAME/KEY: exon
; LOCATION: (163400)...(163586)
; US-10-380-931-24

Query Match          71.2%   Score 17.8; DB 6; Length 169998;
Best Local Similarity 90.5%; Pred. No. 9.6e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AAAAACACTTGGGAAGGGAGAC 22
||||| ||||| ||||| |||||
Db 104105 AAAAAGACTTGGAGGGGAGAC 104125

RESULT 47
US-09-877-177-10
; Sequence 10, Application US/09877177
; Publication No. US20020192652A1
; GENERAL INFORMATION:
; APPLICANT: Peter V. Danenberg et al.
; TITLE OF INVENTION: Method of determining Epidermal Growth
; FILE REFERENCE: 11220/120
; CURRENT APPLICATION NUMBER: US/09/877,177
; CURRENT FILING DATE: 2001-06-11
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 197496
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-877-177-10

Query Match          71.2%   Score 17.8; DB 3; Length 197496;
Best Local Similarity 90.5%; Pred. No. 9.7e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AAAAACACTTGGGAAGGGAGAC 22
||||| ||||| ||||| |||||
Db 112105 AAAAAGACTTGGAGGGGAGAC 112125

RESULT 48
US-10-426-836-10
; Sequence 10, Application US/10426836
; Publication No. US20030211530A1
; GENERAL INFORMATION:
; APPLICANT: K. Danenberg
; TITLE OF INVENTION: Method of determining Epidermal Growth
; FILE REFERENCE: 11220/169
; CURRENT APPLICATION NUMBER: US/10/426,836

```

```

; CURRENT FILING DATE: 2003-05-01
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 197496
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-426-836-10

Query Match          71.2%   Score 17.8; DB 6; Length 197496;
Best Local Similarity 90.5%; Pred. No. 9.7e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AAAAACACTTGGGAAGGGAGAC 22
||||| ||||| ||||| |||||
Db 112105 AAAAAGACTTGGAGGGGAGAC 112125

RESULT 49
US-10-461-862-74
; Sequence 74, Application US/10461862
; Publication No. US20050090434A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc S. Malandro
; TITLE OF INVENTION: Novel Therapeutic Targets in Cancer
; FILE REFERENCE: 529452001800
; CURRENT APPLICATION NUMBER: US/10/461,862
; CURRENT FILING DATE: 2003-06-13
; NUMBER OF SEQ ID NOS: 184
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 74
; LENGTH: 209083
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-461-862-74

Query Match          71.2%   Score 17.8; DB 9; Length 209083;
Best Local Similarity 90.5%; Pred. No. 9.8e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AAAAACACTTGGGAAGGGAGAC 22
||||| ||||| ||||| |||||
Db 112983 AAAAAGACTTGGAGGGGAGAC 113003

RESULT 50
US-10-496-011-1/c
; Sequence 1, Application US/10496011
; Publication No. US20050118588A1
; GENERAL INFORMATION:
; APPLICANT: BOURGERON, THOMAS
; APPLICANT: JAMAIN, STEPHANE
; APPLICANT: QUACH, HELENE
; APPLICANT: BETANCUR, CATALINA
; APPLICANT: LEBROYER, MARION
; APPLICANT: GILLBERG, CHRISTOPHER
; TITLE OF INVENTION: POLYNUCLEOTIDE AND PROTEIN INVOLVED IN SYNAPTOGENESIS, VARIANTS
; FILE REFERENCE: 253820US0XPC7
; CURRENT APPLICATION NUMBER: US/10/496,011
; CURRENT FILING DATE: 2004-05-28
; PRIOR APPLICATION NUMBER: PCT/FR02/04134
; PRIOR FILING DATE: 2002-11-28
; PRIOR APPLICATION NUMBER: CA2364106
; PRIOR FILING DATE: 2001-11-30
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1
; LENGTH: 334462
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:

```

```

; NAME/KEY: exon
; LOCATION: (10670)..(10828)
; FEATURE:
; NAME/KEY: exon
; LOCATION: (83744)..(84513)
; FEATURE:
; NAME/KEY: exon
; LOCATION: (201676)..(201828)
; FEATURE:
; NAME/KEY: exon
; LOCATION: (240359)..(240469)
; FEATURE:
; NAME/KEY: exon
; LOCATION: (312940)..(313125)
; FEATURE:
; NAME/KEY: exon
; LOCATION: (318313)..(319102)
; FEATURE:
; NAME/KEY: exon
; LOCATION: (327018)..(330406)
US-10-496-011-1

```

```

Query Match      71.2%; Score 17.8; DB 9; Length 334462;
Best Local Similarity 90.5%; Pred. No. 1e+03;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 AAAAAACACTTGGAGGGGAGA 21
Db 198779 AAATAACATTGGAGGGGAGA 198759

```

```

Search completed: February 3, 2006, 15:43:41
Job time : 372.556 secs

```


GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 3, 2006, 11:35:31 ; Search time 343.111 Seconds
(without alignments)
61.064 Million cell updates/sec

Title: US-10-719-900-10

Perfect score: 25

Sequence: 1 aaaaacacttggaaggagactca 25

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 6068529 seqs, 419036697 residues

Total number of hits satisfying chosen parameters: 12137058

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database : Published Applications NA New:*

- 1: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	19.2	76.8	600	7	US-10-750-185-2224
C 2	19.2	76.8	600	7	US-10-750-623-2224
C 3	19.2	76.8	777	8	US-11-136-527-3865
C 4	19.2	76.8	777	8	US-11-136-527-7961
C 5	19.2	76.8	3221	7	US-10-750-185-54511
C 6	19.2	76.8	3221	7	US-10-750-623-54511
C 7	18.8	75.2	3481	7	US-10-750-185-27280
C 8	18.8	75.2	3481	7	US-10-750-623-27280
C 9	17.8	71.2	191331	8	US-11-112-908-20
C 10	17.6	70.4	25	8	US-11-136-527-189320
C 11	17.6	70.4	518	8	US-11-136-527-3680
C 12	17.6	70.4	518	8	US-11-136-527-7776
C 13	17.6	70.4	759	7	US-10-750-185-50114
C 14	17.6	70.4	759	7	US-10-750-623-50114
C 15	17.6	70.4	859	8	US-11-136-527-3681
C 16	17.6	70.4	859	8	US-11-136-527-7777
C 17	17.6	70.4	1582	7	US-10-750-185-42140
C 18	17.6	70.4	1582	7	US-10-750-623-42140
C 19	17.6	70.4	1698	7	US-10-750-185-36667
C 20	17.6	70.4	1698	7	US-10-750-623-36667
C 21	17.2	68.8	1805	7	US-10-750-185-53076
C 22	17.2	68.8	1805	7	US-10-750-623-53076

23	17.2	68.8	2655	7	US-10-750-185-35807	Sequence 35807, A
24	17.2	68.8	2655	7	US-10-750-623-35807	Sequence 35807, A
25	17.2	68.8	186442	8	US-11-121-086-104	Sequence 104, App
26	17	68.0	201	7	US-10-995-561-1528	Sequence 1528, Ap
27	17	68.0	201	7	US-10-995-561-1530	Sequence 1530, Ap
28	17	68.0	201	7	US-10-995-561-1538	Sequence 1538, Ap
29	17	68.0	1047	7	US-10-750-185-42937	Sequence 42937, A
30	17	68.0	1047	7	US-10-750-623-42937	Sequence 42937, A
C 31	17	68.0	1613	7	US-10-750-185-62645	Sequence 62645, A
C 32	17	68.0	1613	7	US-10-750-623-62645	Sequence 62645, A
C 33	17	68.0	1698	7	US-10-750-185-56656	Sequence 56656, A
C 34	17	68.0	1698	7	US-10-750-623-56656	Sequence 56656, A
35	17	68.0	1786	7	US-10-750-185-31999	Sequence 31999, A
36	17	68.0	1786	7	US-10-750-623-31999	Sequence 31999, A
37	17	68.0	2038	8	US-11-054-281-29	Sequence 29, Appl
C 38	17	68.0	2494	7	US-10-750-185-27910	Sequence 27910, A
C 39	17	68.0	2494	7	US-10-750-623-27910	Sequence 27910, A
C 40	17	68.0	2583	7	US-10-947-243-67	Sequence 67, Appl
41	17	68.0	2716	7	US-10-995-561-19	Sequence 19, Appl
42	17	68.0	2735	7	US-10-750-185-25095	Sequence 25095, A
43	17	68.0	2735	7	US-10-750-623-25095	Sequence 25095, A
44	17	68.0	3824	7	US-10-750-185-29981	Sequence 29981, A
45	17	68.0	3824	7	US-10-750-623-29981	Sequence 29981, A
C 46	17	68.0	100000	8	US-11-124-368A-2898	Sequence 2898, Ap
47	17	68.0	403278	7	US-10-995-561-13421	Sequence 13421, A
48	17	68.0	1691140	8	US-11-091-018-1	Sequence 1, Appl
C 49	16.8	67.2	82	7	US-10-310-914A-4467	Sequence 4467, Ap
C 50	16.8	67.2	3152	7	US-10-793-626-4056	Sequence 4056, Ap
51	16.8	67.2	160213	8	US-11-121-086-103	Sequence 103, App
C 52	16.6	66.4	25	8	US-11-121-849-81092	Sequence 81092, A
53	16.6	66.4	25	8	US-11-136-527-350443	Sequence 350443, A
54	16.6	66.4	820	7	US-10-750-185-30816	Sequence 30816, A
55	16.6	66.4	820	7	US-10-750-623-30816	Sequence 30816, A
56	16.6	66.4	1400	8	US-11-136-527-8088	Sequence 8088, Ap
57	16.6	66.4	1475	8	US-11-136-527-3992	Sequence 3992, Ap
C 58	16.6	66.4	1789	7	US-10-750-185-64472	Sequence 64472, A
C 59	16.6	66.4	1789	7	US-10-750-623-64472	Sequence 64472, A
C 60	16.6	66.4	1953	7	US-10-750-185-48515	Sequence 48515, A
C 61	16.6	66.4	1953	7	US-10-750-623-48515	Sequence 48515, A
C 62	16.6	66.4	2764	7	US-10-750-185-58029	Sequence 58029, A
C 63	16.6	66.4	2764	7	US-10-750-623-58029	Sequence 58029, A
C 64	16.6	66.4	3660	8	US-11-136-527-3658	Sequence 3658, Ap
C 65	16.6	66.4	215308	7	US-11-121-086-77	Sequence 77, Appl
C 66	16.4	65.6	21	8	US-10-310-914A-963829	Sequence 963829, A
C 67	16.4	65.6	22	7	US-10-310-914A-498609	Sequence 498609, A
C 68	16.4	65.6	201	8	US-11-124-368A-14754	Sequence 14754, A
C 69	16.4	65.6	201	8	US-11-124-368A-14755	Sequence 14755, A
C 70	16.4	65.6	452	5	US-09-978-360A-318	Sequence 318, App
C 71	16.4	65.6	568	5	US-09-978-360A-387	Sequence 387, App
C 72	16.4	65.6	3065	7	US-10-750-185-40303	Sequence 40303, A
C 73	16.4	65.6	3065	7	US-10-750-623-40303	Sequence 40303, A
C 74	16.4	65.6	100000	8	US-11-124-368A-2313	Sequence 2313, Ap
C 75	16.4	65.6	199321	8	US-11-121-086-10	Sequence 10, Appl
C 76	16.2	64.8	25	8	US-11-136-527-350469	Sequence 350469, A
C 77	16.2	64.8	201	8	US-11-124-368A-14496	Sequence 14496, A
C 78	16.2	64.8	986	7	US-10-750-185-53782	Sequence 53782, A
C 79	16.2	64.8	986	7	US-10-750-623-53782	Sequence 53782, A
C 80	16.2	64.8	1141	7	US-10-750-185-36038	Sequence 36038, A
C 81	16.2	64.8	1141	7	US-10-750-623-36038	Sequence 36038, A
C 82	16.2	64.8	2666	7	US-11-090-439-43	Sequence 43, Appl
C 83	16.2	64.8	2666	7	US-10-821-234-215	Sequence 215, App
C 84	16.2	64.8	4162	8	US-11-136-527-2835	Sequence 2835, Ap
C 85	16.2	64.8	38527	8	US-11-124-368A-2312	Sequence 2312, Ap
C 86	16.2	64.8	154548	8	US-11-121-086-33	Sequence 33, Appl
C 87	16.2	64.8	168656	8	US-11-112-908-58	Sequence 58, Appl
C 88	16.2	64.8	170285	8	US-11-112-908-58	Sequence 53, Appl
C 89	16.2	64.8	191343	8	US-11-112-908-53	Sequence 32191, A
C 90	16	64.0	201	7	US-10-995-561-32191	Sequence 51277, A
C 91	16	64.0	201	7	US-10-995-561-51277	Sequence 4266, Ap
C 92	16	64.0	600	8	US-11-136-527-4266	Sequence 4266, Ap
C 93	16	64.0	684	8	US-10-750-185-55780	Sequence 55780, A
C 94	16	64.0	684	8	US-10-750-623-55780	Sequence 55780, A
C 95	16	64.0	756	8	US-11-136-527-170	Sequence 170, App

; LENGTH: 777
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-3865

Query Match 76.8%; Score 19.2; DB 8; Length 777;
Best Local Similarity 87.5%; Pred. No. 10;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAACAACCTTGAAGGAGACTCA 25
||| ||||| ||||| ||||| |||||
DB 567 AAAGACACTTGCAGGGTGACTCA 590

RESULT 4
US-11-136-527-7961
; Sequence 7961, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 7961
; LENGTH: 777
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-7961

Query Match 76.8%; Score 19.2; DB 8; Length 777;
Best Local Similarity 87.5%; Pred. No. 10;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAACAACCTTGAAGGAGACTCA 25
||| ||||| ||||| ||||| |||||
DB 567 AAAGACACTTGCAGGGTGACTCA 590

RESULT 5
US-10-750-185-54511/c
; Sequence 54511, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 54511
; LENGTH: 3221
; TYPE: DNA
; ORGANISM: Bovine 19866880859619
US-10-750-185-54511

Query Match 76.8%; Score 19.2; DB 7; Length 3221;
Best Local Similarity 87.5%; Pred. No. 14;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAACAACCTTGAAGGAGACTCA 25
||| ||||| ||||| ||||| |||||
DB 1164 AAAAACAACCTTGAAGGAGACTCA 1141

RESULT 6
US-10-750-623-54511/c
; Sequence 54511, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 54511
; LENGTH: 3221
; TYPE: DNA
; ORGANISM: Bovine 19866880859619
US-10-750-623-54511

Query Match 76.8%; Score 19.2; DB 7; Length 3221;
Best Local Similarity 87.5%; Pred. No. 14;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAACAACCTTGAAGGAGACTCA 25
||| ||||| ||||| ||||| |||||
DB 1164 AAAAACAACCTTGAAGGAGACTCA 1141

RESULT 7
US-10-750-185-27280/c
; Sequence 27280, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 27280
; LENGTH: 3481
; TYPE: DNA
; ORGANISM: Bovine 19866880859666
US-10-750-185-27280

Query Match 75.2%; Score 18.8; DB 7; Length 3481;
Best Local Similarity 90.9%; Pred. No. 22;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAACAACCTTGAAGGAGACTCA 22
||| ||||| ||||| ||||| |||||

Db 543 AAAAAACACTTAGAAGGAAGAC 522

RESULT 8

US-10-750-623-27280/c
; Sequence 27280, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM11100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 27280
; LENGTH: 3481
; TYPE: DNA
; ORGANISM: Bovine 19866880695666
US-10-750-623-27280

Query Match 75.2%; Score 18.8; DB 7; Length 3481;
Best Local Similarity 90.9%; Pred. No. 22;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAACACTTGAAGGGAGAC 22
|||||
Db 543 AAAAAACACTTAGAAGGAAGAC 522

RESULT 9

US-11-112-908-20
; Sequence 20, Application US/11112908
; Publication No. US20050260659A1
; GENERAL INFORMATION:
; APPLICANT: Harris, Cole
; APPLICANT: Davis, Lisa M.
; TITLE OF INVENTION: Breast Cancer Biomarkers
; FILE REFERENCE: 04-164-US
; CURRENT APPLICATION NUMBER: US/11/112,908
; CURRENT FILING DATE: 2005-04-22
; PRIOR APPLICATION NUMBER: US 60/564,758
; PRIOR FILING DATE: 2004-04-23
; PRIOR APPLICATION NUMBER: US 60/575,978
; PRIOR FILING DATE: 2004-06-01
; PRIOR APPLICATION NUMBER: US 60/631,702
; PRIOR FILING DATE: 2004-11-30
; PRIOR APPLICATION NUMBER: US 60/633,826
; PRIOR FILING DATE: 2004-12-07
; NUMBER OF SEQ ID NOS: 511
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 20
; LENGTH: 191331
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-112-908-20

Query Match 71.2%; Score 17.8; DB 8; Length 191331;
Best Local Similarity 90.5%; Pred. No. 1.5e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAACACTTGAAGGGAGA 21
|||||
Db 36284 AAAAAACCTTGAAGGGGA 36304

RESULT 10

US-11-136-527-189320
; Sequence 189320, Application US/111136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 189320
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Probe
US-11-136-527-189320

Query Match 70.4%; Score 17.6; DB 8; Length 25;
Best Local Similarity 83.3%; Pred. No. 28;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AAAAAACACTTGAAGGGAGACTCA 25
|||||
Db 2 AAAGACACTTGCAGGGGTGACTCA 25

RESULT 11

US-11-136-527-3680
; Sequence 3680, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3680
; LENGTH: 518
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-3680

Query Match 70.4%; Score 17.6; DB 8; Length 518;
Best Local Similarity 83.3%; Pred. No. 53;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AAAAAACACTTGAAGGGAGACTCA 25
|||||
Db 303 AAAGACACTTGCAGGGGTGACTCA 326

RESULT 12

US-11-136-527-7776
; Sequence 7776, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527

; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 7776
; LENGTH: 518
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-7776

Query Match 70.4%; Score 17.6; DB 8; Length 518;
Best Local Similarity 83.3%; Pred. No. 53;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AAAAACACTTGGGAAGGAGACTCA 25
||| ||||| ||||| ||||| |||||
DB 303 AAAGACACTTGCAGGGGTGACTCA 326

RESULT 13
US-10-750-185-50114/c
; Sequence 50114, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 50114
; LENGTH: 759
; TYPE: DNA
; ORGANISM: Bovine 19866881568855
US-10-750-185-50114

Query Match 70.4%; Score 17.6; DB 7; Length 759;
Best Local Similarity 83.3%; Pred. No. 58;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AAAAACACTTGGGAAGGAGACTCA 25
||| ||||| ||||| ||||| |||||
DB 168 AAAAACCTTGGGAAGGATATGGA 145

RESULT 14
US-10-750-623-50114/c
; Sequence 50114, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31

; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 50114
; LENGTH: 759
; TYPE: DNA
; ORGANISM: Bovine 19866881568855
US-10-750-623-50114

Query Match 70.4%; Score 17.6; DB 7; Length 759;
Best Local Similarity 83.3%; Pred. No. 58;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AAAAACACTTGGGAAGGAGACTCA 25
||| ||||| ||||| ||||| |||||
DB 168 AAAAACCTTGGGAAGGATATGGA 145

RESULT 15
US-11-136-527-3681
; Sequence 3681, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3681
; LENGTH: 859
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-3681

Query Match 70.4%; Score 17.6; DB 8; Length 859;
Best Local Similarity 83.3%; Pred. No. 59;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AAAAACACTTGGGAAGGAGACTCA 25
||| ||||| ||||| ||||| |||||
DB 646 AAAGACACTTGCACGGGTGACTCA 669

RESULT 16
US-11-136-527-7777
; Sequence 7777, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 7777
; LENGTH: 859
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-7777

Query Match 70.4%; Score 17.6; DB 8; Length 859;
Best Local Similarity 83.3%; Pred. No. 59;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AAAAACACTTGGGAAGGAGACTCA 25

```
Db      646 AAAGACACTTGCACGGTGACTCA 669
|||||
RESULT 17
US-10-750-185-42140/c
; Sequence 42140, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; PRIOR FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 42140
; LENGTH: 1582
; TYPE: DNA
; ORGANISM: Bovine 19866880823702
US-10-750-185-42140

Query Match      70.4%; Score 17.6; DB 7; Length 1582;
Best Local Similarity 83.3%; Pred. No. 67;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      1 AAAAAACACTTGGAGGGGAGACTC 24
|||||
Db      1249 AAAAAATCCATGTAGGGGAGACTC 1226

RESULT 18
US-10-750-623-42140/c
; Sequence 42140, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 42140
; LENGTH: 1582
; TYPE: DNA
; ORGANISM: Bovine 19866880823702
US-10-750-623-42140

Query Match      70.4%; Score 17.6; DB 7; Length 1582;
Best Local Similarity 83.3%; Pred. No. 67;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      1 AAAAAACACTTGGAGGGGAGACTC 24
|||||
Db      1249 AAAAAATCCATGTAGGGGAGACTC 1226

RESULT 19
US-10-750-185-36667/c
; Sequence 36667, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 36667
; LENGTH: 1698
; TYPE: DNA
; ORGANISM: Bovine 19866880843352
US-10-750-185-36667

Query Match      70.4%; Score 17.6; DB 7; Length 1698;
Best Local Similarity 83.3%; Pred. No. 68;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      2 AAAAAACACTTGGAGGGGAGACTCA 25
|||||
Db      1377 AAAAAATTGAAAGGGATATTCA 1354

RESULT 20
US-10-750-623-36667/c
; Sequence 36667, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 36667
; LENGTH: 1698
; TYPE: DNA
; ORGANISM: Bovine 19866880843352
US-10-750-623-36667

Query Match      70.4%; Score 17.6; DB 7; Length 1698;
Best Local Similarity 83.3%; Pred. No. 68;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      2 AAAAAACACTTGGAGGGGAGACTCA 25
|||||
Db      1377 AAAAAATTGAAAGGGATATTCA 1354

RESULT 21
US-10-750-185-53076
```

; Sequence 53076, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; PRIOR FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 53076
; LENGTH: 1805
; TYPE: DNA
; ORGANISM: Bovine 19866881384060
US-10-750-185-53076

Query Match 68.8%; Score 17.2; DB 7; Length 1805;
Best Local Similarity 86.4%; Pred. No. 1.1e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 AAAACACTTGGAGGAGACTC 24
|||||
DB 992 AAAACACTTAGAAAAGAGACTC 1013

RESULT 22
US-10-750-623-53076
; Sequence 53076, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 53076
; LENGTH: 1805
; TYPE: DNA
; ORGANISM: Bovine 19866881384060
US-10-750-623-53076

Query Match 68.8%; Score 17.2; DB 7; Length 1805;
Best Local Similarity 86.4%; Pred. No. 1.1e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 AAAACACTTGGAGGAGACTC 24
|||||
DB 992 AAAACACTTAGAAAAGAGACTC 1013

RESULT 23
US-10-750-185-35807
; Sequence 35807, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; PRIOR FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 35807
; LENGTH: 2655
; TYPE: DNA
; ORGANISM: Bovine 19866880701412
US-10-750-623-35807

; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 35807
; LENGTH: 2655
; TYPE: DNA
; ORGANISM: Bovine 19866880701412
US-10-750-185-35807

Query Match 68.8%; Score 17.2; DB 7; Length 2655;
Best Local Similarity 86.4%; Pred. No. 1.2e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAACAACCTTGGAGGGAGAC 22
|||||
DB 2257 AACAAACACCTGGAAGTGAGAC 2278

RESULT 24
US-10-750-623-35807
; Sequence 35807, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 35807
; LENGTH: 2655
; TYPE: DNA
; ORGANISM: Bovine 19866880701412
US-10-750-623-35807

Query Match 68.8%; Score 17.2; DB 7; Length 2655;
Best Local Similarity 86.4%; Pred. No. 1.2e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAACAACCTTGGAGGGAGAC 22
|||||
DB 2257 AACAAACACCTGGAAGTGAGAC 2278

RESULT 25
US-11-121-086-104
; Sequence 104, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; APPLICANT: NIELSEN, KIRSTEN V.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES

```
; FILE REFERENCE: 09138.6000-00000
; CURRENT APPLICATION NUMBER: US/11/121,086
; CURRENT FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 104
; LENGTH: 186442
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-104

Query Match      68.8%; Score 17.2; DB 8; Length 186442;
Best Local Similarity 86.4%; Pred. No. 2.7e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AAAAAACACTTGGAGGGAGAC 22
    ||||| ||||| ||||| |||||
Db 128241 AATATACACTTGGAGAGAGCC 128262

RESULT 26
US-10-995-561-1528
; Sequence 1528, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1528
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-1528

Query Match      68.0%; Score 17; DB 7; Length 201;
Best Local Similarity 80.0%; Pred. No. 83;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 AAAAAACACTTGGAGGGAGACTCA 25
    ||||| ||||| ||||| ||||| |||||
Db 132 AATAGATGCTTGTAGGGAGATTCA 156

RESULT 27
US-10-995-561-1530
; Sequence 1530, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1530
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-1530

Query Match      68.0%; Score 17; DB 7; Length 201;
Best Local Similarity 80.0%; Pred. No. 83;

Qy 1 AAAAAACACTTGGAGGGAGACTCA 25
    ||||| ||||| ||||| ||||| |||||
Db 132 AATAGATGCTTGTAGGGAGATTCA 156

RESULT 28
US-10-995-561-1538
; Sequence 1538, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1538
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-1538

Query Match      68.0%; Score 17; DB 7; Length 201;
Best Local Similarity 80.0%; Pred. No. 83;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 AAAAAACACTTGGAGGGAGACTCA 25
    ||||| ||||| ||||| ||||| |||||
Db 168 AATAGATGCTTGTAGGGAGATTCA 192

RESULT 29
US-10-750-185-42937
; Sequence 42937, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERE, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 42937
; LENGTH: 1047
; TYPE: DNA
; ORGANISM: Bovine
US-10-750-185-42937

Query Match      68.0%; Score 17; DB 7; Length 1047;
Best Local Similarity 80.0%; Pred. No. 1.2e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 AAAAAACACTTGGAGGGAGACTCA 25
    ||||| ||||| ||||| ||||| |||||
Db 89 AATAAAACATGGAAGGGAGACTCA 113

RESULT 30
US-10-750-623-42937
```



```
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; PRIOR FILING DATE: 2003-12-31
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 56656
; LENGTH: 1698
; TYPE: DNA
; ORGANISM: Bovine 19866880612383
US-10-750-623-56656

Query Match      68.0%; Score 17; DB 7; Length 1698;
Best Local Similarity 80.0%; Pred. No. 1.3e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAAACACTTGGGAAGGGGAGACTCA 25
   ||||| ||||| ||||| ||||| |||||
Db 783 AAAAAACACTTAGAAGAGTGCCTGA 759

RESULT 35
US-10-750-185-31999
; Sequence 31999, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 31999
; LENGTH: 1786
; TYPE: DNA
; ORGANISM: Bovine 19866880848133
US-10-750-185-31999

Query Match      68.0%; Score 17; DB 7; Length 1786;
Best Local Similarity 80.0%; Pred. No. 1.3e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAAACACTTGGGAAGGGGAGACTCA 25
   ||||| ||||| ||||| ||||| |||||
Db 681 AATAACAGATGGAAGAGAGACTGA 705

RESULT 36
US-10-750-623-31999
; Sequence 31999, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
```

```
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; PRIOR FILING DATE: 2003-12-31
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 31999
; LENGTH: 1786
; TYPE: DNA
; ORGANISM: Bovine 19866880848133
US-10-750-623-31999

Query Match      68.0%; Score 17; DB 7; Length 1786;
Best Local Similarity 80.0%; Pred. No. 1.3e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAAACACTTGGGAAGGGGAGACTCA 25
   ||||| ||||| ||||| ||||| |||||
Db 681 AATAACAGATGGAAGAGAGACTGA 705

RESULT 37
US-11-054-281-29
; Sequence 29, Application US/11054281
; Publication No. US20060013813A1
; GENERAL INFORMATION:
; APPLICANT: Mezes et al.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-240CIP
; CURRENT APPLICATION NUMBER: US/11/054,281
; CURRENT FILING DATE: 2005-02-08
; PRIOR APPLICATION NUMBER: 60/261,014
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/261,018
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/318,410
; PRIOR FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/261,013
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/261,026
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/261,029
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/313,170
; PRIOR FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: 10/044,564
; PRIOR FILING DATE: 2002-01-11
; NUMBER OF SEQ ID NOS: 324
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 29
; LENGTH: 2038
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-054-281-29

Query Match      68.0%; Score 17; DB 8; Length 2038;
Best Local Similarity 80.0%; Pred. No. 1.4e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAAACACTTGGGAAGGGGAGACTCA 25
   ||||| ||||| ||||| ||||| |||||
Db 1719 AAAAGATGCTTGTGAAGGGGAGATTCA 1743

RESULT 38
US-10-750-185-27910/c
; Sequence 27910, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
```

; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 27910
; LENGTH: 2494
; TYPE: DNA
; ORGANISM: Bovine 19866880779532
US-10-750-185-27910

Query Match 68.0%; Score 17; DB 7; Length 2494;
Best Local Similarity 80.0%; Pred. No. 1.4e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 AAAAAACACTTGGAGGGAGACTCA 25
||||| ||||| ||||| ||||| |||||
Db 1477 AAAAAGCATTTTGAAGTGACACTCA 1453

RESULT 39
US-10-750-623-27910/c
; Sequence 27910, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 27910
; LENGTH: 2494
; TYPE: DNA
; ORGANISM: Bovine 19866880779532
US-10-750-623-27910

Query Match 68.0%; Score 17; DB 7; Length 2494;
Best Local Similarity 80.0%; Pred. No. 1.4e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 AAAAAACACTTGGAGGGAGACTCA 25
||||| ||||| ||||| ||||| |||||
Db 1477 AAAAAGCATTTTGAAGTGACACTCA 1453

RESULT 40
US-10-947-249-67/c
; Sequence 67, Application US/10947249
; Publication No. US20050287541A1
; GENERAL INFORMATION:
; APPLICANT: Akita NAKAGAWARA
; APPLICANT: Miki OHIRA
; APPLICANT: Shin ISHII
; APPLICANT: Takeshi GOTO

; APPLICANT: HIROYUKI KUBO
; APPLICANT: Takahiro HIRATA
; APPLICANT: Yasuko YOSHIDA
; APPLICANT: Saichi YAMADA
; TITLE OF INVENTION: Microarray for Predicting the Prognosis of Neuroblastoma
; FILE REFERENCE: 117007
; CURRENT APPLICATION NUMBER: US/10/947,249
; CURRENT FILING DATE: 2004-09-23
; PRIOR APPLICATION NUMBER: US 60/505,614
; PRIOR APPLICATION NUMBER: 2003-09-25
; NUMBER OF SEQ ID NOS: 200
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 67
; LENGTH: 2583
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-947-249-67

Query Match 68.0%; Score 17; DB 7; Length 2583;
Best Local Similarity 80.0%; Pred. No. 1.4e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 AAAAAACACTTGGAGGGAGACTCA 25
||||| ||||| ||||| ||||| |||||
Db 1019 AAAAAAGCTTGGATGCGGCTCA 995

RESULT 41
US-10-995-561-19
; Sequence 19, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 2716
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-19

Query Match 68.0%; Score 17; DB 7; Length 2716;
Best Local Similarity 80.0%; Pred. No. 1.4e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 AAAAAACACTTGGAGGGAGACTCA 25
||||| ||||| ||||| ||||| |||||
Db 2240 AAAAGATGCTTGTAAAGGGAGATTCA 2264

RESULT 42
US-10-750-185-25095
; Sequence 25095, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31


```

; LENGTH: 1691140
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(1691140)
; OTHER INFORMATION: n=A,T,C or G
US-11-091-018-1

Query Match      68.0%; Score 17; DB 8; Length 1691140;
Best Local Similarity 80.0%; Pred. No. 3.3e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY      1 AAAAAACACTTGGAGGGGAGACTCA 25
      | | | | | | | | | | | | | | | |
Db      1201842 ACAATACACTTGAATGGTGAGTCA 1201866

RESULT 49
US-10-310-914A-4467/c
; Sequence 4467, Application US/10310914A
; Publication No. US20060003322A1
; GENERAL INFORMATION:
; APPLICANT: Bentwich, Isaac
; APPLICANT: Shiler, Knyzat
; TITLE OF INVENTION: Bioinformatically detectable group of novel regulatory g
; FILE REFERENCE: 06087.0200.CPUS01
; CURRENT APPLICATION NUMBER: US/10/310,914A
; CURRENT FILING DATE: 2002-12-06
; NUMBER OF SEQ ID NOS: 1388402
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 4467
; LENGTH: 82
; TYPE: RNA
; ORGANISM: Human
US-10-310-914A-4467

Query Match      67.2%; Score 16.8; DB 7; Length 82;
Best Local Similarity 90.0%; Pred. No. 85;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 AAAAAACACTTGGAGGGGAG 20
      | | | | | | | | | | | | | |
Db      63 AAAAAACACTTGGAGGTGGG 44

RESULT 50
US-10-793-626-4056/c
; Sequence 4056, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4056
; LENGTH: 3152
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-4056

Query Match      67.2%; Score 16.8; DB 7; Length 3152;
Best Local Similarity 90.0%; Pred. No. 1.8e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 1 AAAAAACACTTGGAGGGAG 20
| | | | | | | | | | | | | | | |
Db 1182 AAAAAACATTTCGAGGGAG 1163

Search completed: February 3, 2006, 16:19:56
Job time : 347.111 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 3, 2006, 10:53:23 ; Search time 863 Seconds
(without alignments)
1646.682 Million cell updates/sec

Title: US-10-719-900-15

Perfect score: 25
Sequence: 1 aaaaacacacacagcgagcctcgg 25

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues
Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 150 summaries

Database :

GenEmbl:*
1: gb_ba:*
2: gb_in:*
3: gb_env:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pr:*
9: gb_ro:*
10: gb_sts:*
11: gb_sy:*
12: gb_un:*
13: gb_vi:*
14: gb_htg:*
15: gb_pl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	23.4	93.6	2331	9 BC005588	BC005588 Mus muscu
2	23.4	93.6	17372	9 AC117831	AC117831 Mus muscu
3	23.4	93.6	205272	9 AC118646	AC118646 Mus muscu
4	23.4	93.6	246494	14 AC120454	AC120454 Rattus no
5	19.2	76.8	110000	2 CP000079_06	Continuation (7 of
6	19.2	76.8	127380	2 AC113585	AC113585 Trypanoso
7	18.8	75.2	2426	5 BC071044	BC071044 Xenopus 1
8	18.8	75.2	149754	9 AC158827	AC158827 Mus muscu
9	18.8	75.2	163916	14 AC148067	AC148067 Canis fam
10	18.6	74.4	791	5 CR386415	CR386415 Gallus ga
11	18.6	74.4	10029	1 AB013307	AB013307 Methanosa
12	18.6	74.4	54569	9 BX571685	BX571685 Mouse DNA
13	18.6	74.4	65758	14 AC100392	AC100392 Mus muscu
14	18.6	74.4	97626	9 AL845429	AL845429 Mouse DNA
15	18.6	74.4	109295	14 AC141049	AC141049 Rattus no
16	18.6	74.4	110000	14 AC119467_1	Continuation (2 of
17	18.6	74.4	110000	15 AP008216_225	Continuation (226
18	18.6	74.4	139999	15 AC018727	AC018727 Oryza sat

c	19	18.6	74.4	156342	9	AC158566	AC158566 Mus muscu
	20	18.6	74.4	157085	14	AC160549	AC160549 Mus muscu
	21	18.6	74.4	192561	9	AL805943	AL805943 Mouse DNA
	22	18.6	74.4	196097	8	AC122175	AC122175 Pan trogl
	23	18.6	74.4	214864	14	AC162627	AC162627 Bos tauru
	24	18.6	74.4	218219	14	AC154992	AC154992 Bos tauru
	25	18.6	74.4	218746	9	AC115721	AC115721 Mus muscu
	26	18.6	74.4	230451	14	AC138614	AC138614 Mus muscu
c	27	18.6	74.4	236746	14	AC119312	AC119312 Rattus no
	28	18.6	74.4	258567	14	CR376791	CR376791 Mus muscu
	29	18.6	74.4	300029	15	AB017122	AB017122 Oryza sat
	30	18.4	73.6	142359	14	EX322621	EX322621 Danio rer
	31	18.4	73.6	158420	14	CR936300	CR936300 Danio rer
c	32	18.2	72.8	252	15	CUSCC	D63386 Cumele sat
	33	18.2	72.8	1181	8	HS323695	AJ323695 Homo sapi
	34	18.2	72.8	1824	13	AY374526	AY374526 Borna dis
	35	18.2	72.8	1824	13	AY374527	AY374527 Borna dis
c	36	18.2	72.8	10029	1	AE011319	AE011319 Leptospi
	37	18.2	72.8	37857	14	AC142512	AC142512 Rattus no
c	38	18.2	72.8	148875	9	AC101757	AC101757 Mus muscu
	39	18.2	72.8	150313	8	AC137672	AC137672 Homo sapi
	40	18.2	72.8	156107	8	AC099543	AC099543 Homo sapi
c	41	18.2	72.8	190122	14	AC083820	AC083820 Rattus no
	42	18.2	72.8	212822	14	AC080156	AC080156 Rattus no
	43	18.2	72.8	225030	9	AC107608	AC107608 Rattus no
c	44	18.2	72.8	247137	14	AC098268	AC098268 Rattus no
	45	18.2	72.8	249741	14	AC096330	AC096330 Rattus no
c	46	18.2	72.8	271995	14	AC080155	AC080155 Rattus no
	47	18.2	72.8	301258	1	AE017296	AE017296 Leptospi
	48	18.2	72.8	349751	1	CR378680	CR378680 Photobact
c	49	17.8	71.2	1368	6	AX432478	AX432478 Sequence
	50	17.8	71.2	2347	6	AX954557	AX954557 Sequence
c	51	17.8	71.2	110000	1	AE017333_01	Continuation (2 of
	52	17.8	71.2	110000	1	CP000002_01	Continuation (2 of
c	53	17.8	71.2	115780	6	CQ869701	CQ869701 Sequence
	54	17.8	71.2	178882	14	AC147494	AC147494 Carolia
	55	17.8	71.2	178882	14	AC158270	AC158270 Carolia
	56	17.8	71.2	208684	9	AC122272	AC122272 Mus muscu
	57	17.8	71.2	215602	9	AC155330	AC155330 Mus muscu
	58	17.8	71.2	225535	14	AC111381	AC111381 Rattus no
	59	17.6	70.4	559	6	AX438410	AX438410 Sequence
	60	17.6	70.4	597	6	AX438347	AX438347 Sequence
	61	17.6	70.4	2018	15	MGU12335	U12335 Magnaporthe
	62	17.6	70.4	6073	6	CQ490795	CQ490795 Sequence
	63	17.6	70.4	6073	6	CQ496641	CQ496641 Sequence
c	64	17.6	70.4	53675	8	AL391708	AL391708 Human DNA
c	65	17.6	70.4	69058	14	AC165884	AC165884 Bos tauru
	66	17.6	70.4	69358	14	AC164902	AC164902 Bos tauru
c	67	17.6	70.4	94695	15	ATF23E13	AL022141 Arabidops
	68	17.6	70.4	110000	1	AE016822_05	Continuation (6 of
	69	17.6	70.4	110000	1	AY596297_06	Continuation (7 of
c	70	17.6	70.4	110000	1	AY596297_07	Continuation (8 of
	71	17.6	70.4	110000	1	AY596297_07	Continuation (8 of
	72	17.6	70.4	110000	1	CP000058_40	Continuation (41 o
c	73	17.6	70.4	110000	14	AP008212_080	Continuation (3 o
	74	17.6	70.4	110000	15	AP008212_080	Continuation (81 o
	75	17.6	70.4	114691	4	AC144636	AC144636 Artibeus
	76	17.6	70.4	117036	5	CR847979	CR847979 Zebrafish
	77	17.6	70.4	129135	8	AL136162	AL136162 Human DNA
	78	17.6	70.4	145342	14	AC023320	AC023320 Homo sapi
	79	17.6	70.4	153734	15	AP003490	AP003490 Oryza sat
	80	17.6	70.4	155544	15	AP005518	AP005518 Oryza sat
	81	17.6	70.4	158431	14	AC147528	AC147528 Oryza sat
c	82	17.6	70.4	161023	14	AC069508	AC069508 Homo sapi
	83	17.6	70.4	172894	14	AC150571	AC150571 Bos tauru
	84	17.6	70.4	172963	7	AY962392	AY962392 Aeromonas
c	85	17.6	70.4	173591	7	AY375531	AY375531 Bacteriop
c	86	17.6	70.4	177640	8	AC079456	AC079456 Homo sapi
	87	17.6	70.4	178079	14	AC068929	AC068929 Homo sapi
	88	17.6	70.4	180223	8	AC062017	AC062017 Homo sapi
	89	17.6	70.4	180727	9	AC140278	AC140278 Mus muscu
	90	17.6	70.4	183008	14	AC019276	AC019276 Homo sapi
	91	17.6	70.4	188255	14	AC074241	AC074241 Homo sapi

92	17.6	70.4	188921	8	AC096730	AC096730 Homo sapi	ORGANISM	Mus musculus
93	17.6	70.4	189385	8	AL592287	AL592287 Human DNA	REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridea; Muridae; Mus.
94	17.6	70.4	191281	14	AC013624	AC013624 Homo sapi	AUTHORS	1 (bases 1 to 2331)
95	17.6	70.4	191755	8	AC021016	AC021016 Homo sapi		Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, P.S., Wagner, L.H., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Datchenko, L., Marusina, K., Farmer, A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheet, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettner, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, Y.S., Krzywinski, M.I., Skalska, U., Smallus, D.E., Sutterfield, A., Schein, J.E., Jones, S.J. and Marra, M.A.
96	17.6	70.4	192865	5	AL845282	AL845282 Zebrafish		Mammalian Gene Collection Program Team
97	17.6	70.4	196533	8	CNS000VYI	CNS000VYI		Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
98	17.6	70.4	198750	15	ATCHRIV85	ATCHRIV85		Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
99	17.6	70.4	203056	5	BX346480	BX346480 Zebrafish		12477932
100	17.6	70.4	205265	14	AC146621	AC146621 Otofemur		2 (bases 1 to 2331)
101	17.6	70.4	209881	14	AC150988	AC150988 Bos taurus		NIH MGC Project
102	17.6	70.4	229484	14	AC094760	AC094760 Rattus no		Direct Submission
103	17.6	70.4	232278	14	AC087876	AC087876 Bos taurus		Submitted (02-APR-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Bethesda, MD 20892-2590, USA
104	17.6	70.4	240531	14	AC149658	AC149658 Bos taurus		NIH-MGC Project URL: http://mgc.nci.nih.gov
105	17.6	70.4	250579	14	AC156053	AC156053 Bos taurus		Contact: MGC help desk
106	17.4	69.6	87637	8	AL390960	AL390960 Human DNA		Email: cgapbs-r@mail.nih.gov
107	17.4	69.6	110000	1	CP000091_18	Continuation (19 o		Tissue Procurement: Jeffrey Green M.D.
108	17.4	69.6	174086	5	CR759824	CR759824 Zebrafish		CDNA Library Preparation: Life Technologies, Inc.
109	17.4	69.6	261493	14	AC106564	AC106564 Rattus no		CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
110	17.2	68.8	778	10	BV597660	BV597660 S215P6141		DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center
111	17.2	68.8	1803	15	BV673357	BV673357 Peziza ar		Center code: BCM-HGSC
112	17.2	68.8	1849	5	BC086711	BC086711 Panio rer		Web site: http://www.hgsc.bcm.tmc.edu/cdna/
113	17.2	68.8	1968	2	AY730404	AY730404 Daphnia l		Contact: angbcm.tmc.edu
115	17.2	68.8	2352	8	AB097523	AB097523 Macaca fa		Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati, A.N., Gibbs, R.A.
116	17.2	68.8	56154	14	AC079961	AC079961 Homo sapi		Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
117	17.2	68.8	61450	8	HS15D7	AL031229 Human DNA		Series: IRAK Plate: 17 Row: m Column: 23
118	17.2	68.8	62315	8	AC106802	AC106802 Homo sapi		This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, similarity but not identity to protein.
119	17.2	68.8	75458	14	CR954203_09	Continuation (16 o		Location/Qualifiers
120	17.2	68.8	81885	9	AL663040	AL663040 Mouse DNA		1. .2331
121	17.2	68.8	94874	9	AL808126	AL808126 Mouse DNA		/organism="Mus musculus"
122	17.2	68.8	96614	8	HS136B1	AL031768 Human DNA		/mol_type="mRNA"
123	17.2	68.8	107439	15	AC148527	AC148527 Medicago		/strain="FVB/N"
124	17.2	68.8	110000	1	CP000023_15	Continuation (16 o		/db_xref="taxon:10090"
125	17.2	68.8	110000	1	CP000024_15	Continuation (16 o		/clone="MGC:11579 IMAGE:3708463"
126	17.2	68.8	110000	14	AC109710_0	AC109710 Rattus no		/tissue_type="Mammary tumor, C3(1)-Tag model. Infiltrating ductal carcinoma. 5 month old virgin mouse."
127	17.2	68.8	110000	14	CR954211_2	Continuation (3 of		/clone_lib="NCI CGAP_Mam6"
128	17.2	68.8	110000	14	CR954211_3	Continuation (4 of		/lab_host="DH10B"
129	17.2	68.8	110000	14	CT005251_2	Continuation (3 of		/note="Vector: pCMV-SPORT6"
130	17.2	68.8	110000	15	CR380958_04	Continuation (5 of		1. .2331
131	17.2	68.8	110000	15	AP008209_055	Continuation (56 o		/gene="Tegt"
132	17.2	68.8	110000	15	AP008210_062	Continuation (63 o		/db_xref="GeneID:110213"
133	17.2	68.8	110000	15	AP008210_253	Continuation (254		/db_xref="MGI:99682"
134	17.2	68.8	110000	15	AP008212_015	Continuation (16 o		68. .781
135	17.2	68.8	110000	15	AP008212_016	Continuation (17 o		/gene="Tegt"
136	17.2	68.8	111686	15	F14L17	AC012188 Sequence		
137	17.2	68.8	119716	14	AC147012	AC147012 Medicago		
138	17.2	68.8	124028	8	AC022417	AC022417 Homo sapi		
139	17.2	68.8	143113	15	OSJN00559	AP000559 Oryza sat		
140	17.2	68.8	147108	15	OSJN00115	AL606397 Oryza sat		
141	17.2	68.8	149042	14	AC016410	AL016410 Homo sapi		
142	17.2	68.8	152990	14	AC130432	AC130432 Homo sapi		
143	17.2	68.8	153126	15	AC132214	AC132214 Genomic s		
144	17.2	68.8	154089	5	BX649302	BX649302 Zebrafish		
145	17.2	68.8	155643	15	AP000391	AP000391 Oryza sat		
146	17.2	68.8	156244	15	CNS0707IX	AL627362 DNA centr		
147	17.2	68.8	161940	14	AC148232	AC148232 Didelphis		
148	17.2	68.8	164577	15	OSJN00169	AL662969 Oryza sat		
149	17.2	68.8	168031	15	AC105928	AC105928 Oryza sat		
150	17.2	68.8	170401	14	AC165385	AC165385 Oryctolag		

RESULT 1

BC005588

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

BC005588

Mus musculus testis enhanced gene transcript, mRNA (cdna clone MGC:11579 IMAGE:3708463), complete cds.

BC005588

BC005588.1

GI:13542767

Mus musculus (house mouse)

gene

CDS


```

/codon_start=1
/product="testis enhanced gene transcript"
/protein_id="AA05588.1"
/db_xref="GI:13542768"
/db_xref="GeneID:110213"
/db_xref="WGI:99682"
/translations="MWIFKINFDAALFKFSHTPTQOHLKKVAFALCMFVAAAG
AYVVTHTFIQAGLSALMIWLMATPHSHETQKRLGLLAGFAFTVGLGPGA
LELCIAVNSILPTAMGTAMITFCFSLAYARRSYLFLGILMSANSLMLLSLG
NLFPGSIWUFOANLYIGLLVMCGFVLFDQILITKRAHGDKDYHHCVDLFLDFTLFL
RKLMLILAFNEKDKKKEK"

```

ORIGIN

```

Query Match      93.6%  Score 23.4;  DB 9;  Length 2331;
Best Local Similarity 96.0%  Pred. No. 3.4;
Matches 24;  Conservative 0;  Mismatches 1;  Indels 0;  Gaps 0;

```

```
Qy 1 AAAAAACGACACGAGCGACCTCGG 25
|||||

```

```
Db 2008 AAAAAACGACACTCGGACCTCGG 2032
|||||

```

RESULT 2

```
AC117831 173772 bp DNA linear ROD 17-FEB-2005
Mus musculus chromosome 15, clone RP24-106E17, complete sequence.
AC117831
AC117831.6 GI:59891528
HTG.

```

SOURCE

```
Mus musculus (house mouse)

```

```
Mus musculus

```

```
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 173772)

```

```
Birren,B., Nussbaum,C. and Lander,E.

```

```
Mus musculus chromosome 15, clone RP24-106E17

```

```
Unpublished

```

```
2 (bases 1 to 173772)

```

```
Birren,B., Linton,L., Nussbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L.,
Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,
Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faroo,S., Ferreira,P., FitzGerald,M., Gage,D., Gage,J., Gardyna,S.,
Gande,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N., Jones,C.,
Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
Lander,T., Lehoczy,J., Levine,R., Lindblad-Toh,K., Liu,G.,
MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L.,
Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,
Norbu,C., Norman,K., O'Connor,T., O'Donnell,P., O'Neill,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Rilev,R., Rise,C., Rogov,P.,
Roman,J., Roberti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Traversman,A., Talamas,J., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

```

```
Direct Submission

```

```
Submitted (10-APR-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA

```

```
3 (bases 1 to 173772)

```

```
Birren,B., Nussbaum,C., Lander,E., Abouelleil,A., Allen,N.,
Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V.,
Bloom,T., Boguslavsky,L., Boukhgalter,B., Camarata,J., Chang,J.,
Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B.,
Dearellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L.,
Erickson,J., Faro,S., Ferreira,P., FitzGerald,M., Gage,D.,
Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N.,
Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I.,

```

TITLE

JOURNAL

REFERENCE

AUTHORS

FEATURES

source

```

1. 173772
   Location/Qualifiers
     organism="Mus musculus"
     mol_type="genomic DNA"
     db_xref="taxon:10090"
     chromosome="15"
     map="15"
     clone="RP24-106E17"
     clone_lib="RPCI-24 Male Mouse BAC"
     repeat_region
       859..1005
       /rpt_family="B1_MM"
     repeat_region
       1021..1197
       /rpt_family="B3"
     repeat_region
       1440..1550
       /rpt_family="RSINE1"
     repeat_region
       2148..2168

```

Some of the sequence contained within base pairs 119388 to the end of the clone was stolen from accession AC139317.

Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbitt,R., MacLean,C., Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nguyen,T., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Traversman,A., Talamas,J., Trigilio,J., Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission

Submitted (20-JAN-2005) Broad Institute of MIT and Harvard, 320

Charles Street, Cambridge, MA 02141, USA

4 (bases 1 to 173772)

REFERENCE

AUTHORS

Birren,B., Nussbaum,C., Lander,E., Abouelleil,A., Allen,N., Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B., Dearellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbitt,R., MacLean,C., Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nguyen,T., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Traversman,A., Talamas,J., Trigilio,J., Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission

Submitted (17-FEB-2005) Broad Institute of MIT and Harvard, 320

Charles Street, Cambridge, MA 02141, USA

On Feb 17, 2005 this sequence version replaced gi:57977528.

All repeats were identified using RepeatMasker:

Smith, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Broad Institute of MIT and Harvard

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence_submissions@broad.mit.edu

----- Project Information

Center project name: L25872

Center clone name: 106_E_17

```
/rpt family="AT_rich"
2412..2435
/rpt family="(TTTG)n"
2811..2942
/rpt family="B1_MM"
2943..2979
/rpt family="(CA)n"
complement(2985..3100)
/rpt family="B2_Mm2"
complement(3186..3678)
/rpt family="PB1D9"
complement(3690..3894)
/rpt family="MER82"
4129..4187
/rpt family="(TG)n"
5087..5112
/rpt family="AT_rich"
5203..5290
/rpt family="B1_MM"
complement(6980..7144)
/rpt family="ID_B1"
complement(8010..8148)
/rpt family="B1_MM"
complement(8160..8341)
/rpt family="B3"
complement(10674..10867)
/rpt family="ORR1D"
complement(11222..11321)
/rpt family="RSINE1"
complement(12570..12717)
/rpt family="B1_MM"
complement(12967..13089)
/rpt family="B3"
complement(13073..13201)
/rpt family="ID_B1"
13265..13298
/rpt family="(CG)n"
complement(14043..14228)
14448..14641
/rpt family="B3A"
14654..14788
/rpt family="B1_MM"
14790..14816
/rpt family="(TAAAA)n"
complement(14948..15137)
15188..15204
/rpt family="B2_Mm1"
/rpt family="AT_rich"
15205..15280
/rpt family="ID4"
15281..15311
/rpt family="AT_rich"
15312..15430
/rpt family="B1F"
15427..15501
/rpt family="ID_B1"
15559..15667
/rpt family="(TA)n"
15670..15783
/rpt family="L1MC/D"
complement(15799..15857)
/rpt family="B3A"
complement(15828..15857)
15858..15899
/rpt family="(TTTTC)n"
15903..15961
/rpt family="(TC)n"
complement(15962..16108)
/rpt family="B2_Mm2"
complement(16129..16356)
/rpt family="ID_B1"

repeat_region 17030..17068
/rpt family="T-rich"
repeat_region 17132..17191
/rpt family="PB1D7"
repeat_region 17192..17220

Query Match 93.6%; Score 23.4; DB 9; Length 173772;
Best Local Similarity 96.0%; Pred. No. 2.8;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAAAACGACAGCGAGCCCTCGG 25
|||||
Db 9971 AAAAAACGACACTGCGAGCCCTCGG 9995
|||||

RESULT 3
AC118646 205272 bp DNA linear ROD 25-DEC-2004
LOCUS Mus musculus chromosome 15, clone RP24-116A18, complete sequence.
DEFINITION AC118646
ACCESSION AC118646
VERSION AC118646.23 GI:56799106
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Muridae; Murinae; Mus.

REFERENCE
1 (bases 1 to 205272)
Birren,B., Nusbaum,C. and Lander,B.
Mus musculus chromosome 15, clone RP24-116A18
Unpublished
2 (bases 1 to 205272)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L.,
Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
Chazaro,B., Choepel,Y., Collangelo,M., Collins,S., Collymore,A.,
Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., Lakocque,K., Lamazares,R.,
Lander,T., Lehoczy,J., Levine,R., Lindblad-Toh,K., Liu,G.,
MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., Meldrum,J., Meneus,L.,
Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (19-APR-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 205272)
Birren,B., Nusbaum,C., Lander,B., Abouelleil,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L.,
Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
Chazaro,B., Choepel,Y., Collangelo,M., Collins,S., Collymore,A.,
Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Dooley,K., Dorris,L.,
Drickson,J., Faro,S., Ferreira,P., FitzGerald,M., Gage,D.,
Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N.,
Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I.,
Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Lander,T.,
Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbitt,R.,
MacLean,C., Macdonald,P., Major,J., Manning,J., Matthews,C.,
McCarthy,M., Meldrum,J., Meneus,L., Mihova,T., Mlenga,V.,
Murphy,T., Naylor,J., Nguyen,C., Nguyen,T., Nicol,R., Norbu,C.,
O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C.,
```

Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE

Submitted (06-NOV-2004) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE

Birren, B., Nusbbaum, C., Lander, E., Abouelleil, A., Allen, N., Anderson, S., Anderson, S., Arachchi, H. M., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhalgalier, B., Camarata, J. C., Chang, J., Choepel, Y., Collamore, A., Cooke, A., Cooke, P., Corum, B., DeArellano, K., Diaz, J. S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Galdyna, S., Graham, L., Grand-pierre, N., Hafez, N., Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kellis, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Liu, X., Lui, A., Mabbitt, R., McLean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrum, J., Meneus, L., Mihova, T., Milenga, V., Murphy, T., Naylor, K., Nguyen, C., Nguyen, T., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramsamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schuback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J., Topfkan, K., Travers, M., Vassiliev, H., Venkataraman, V. S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE

Submitted (25-DEC-2004) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Dec 25, 2004 this sequence version replaced gi:55468782.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

COMMENT

Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center
Center: Whitehead Institute/MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

web site: <http://www.seq.wi.mit.edu>
Contact: sequence_submissions@broad.mit.edu

----- Project Information

Center project name: L25047

Center clone name: 116 A 18

FEATURES

	Location/Qualifiers		repeat_region	/rpt_family="B1_MM"
TUES	source		repeat_region	21435. .21477
	1. .205272		repeat_region	/rpt_family="(TCTCTG)n"
	/organism="Mus musculus"		repeat_region	complement(21484. .21592)
	/mol_type="genomic DNA"		unsure	/rpt_family="B1_MM"
	/db_xref="taxon:10090"			complement(22963. .23036)
	/chromosome="15"			/note="single clone coverage"
	/map="15"		repeat_region	24323. .24427
	/clone="RP24-116A18"		repeat_region	/rpt_family="B1F"
	/clone_lib="RPCI-24 Male Mouse BAC"		repeat_region	26361. .26496
	1692. .1786		repeat_region	/rpt_family="MTD"
	/rpt_family="MIR3"		repeat_region	26496. .26566
	2376. .2415		repeat_region	/rpt_family="MTD"
	/rpt_family="MIR"		repeat_region	26567. .26759
	3053. .3097		repeat_region	/rpt_family="B2_Mm1"
	/rpt_family="(TG)n"		repeat_region	26760. .26847
	3369. .3443		repeat_region	/rpt_family="MTD"
	/rpt_family="CT-rich"		repeat_region	26931. .26953
	4222. .4368		repeat_region	/rpt_family="(TTTG)n"
	/rpt_family="RSINE1"		repeat_region	complement(27072. .27412)
	4552. .4732		repeat_region	/rpt_family="MTE"
	/rpt_family="C-rich"		repeat_region	complement(27529. .27656)
	5224. .5370			/rpt_family="B1_MM"
	/rpt_family="B1_MM"			
	5372. .5393			
	/rpt_family="(CAAA)n"			
	6268. .6314			

Query Match 93.6%; Score 23.4; DB

Query Match

93.6%; Score 23.4; DB 9; Length 205272;

Best Local Similarity 96.0%; Pred. No. 2,7; Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AAAAAACGACACGAGCCCTCGG 25
|||||
Db 183211 AAAAAACGACACTGGCGCCTCGG 181235

RESULT 4
AC120454 246494 bp DNA linear HTG 15-NOV-2002
LOCUS Rattus norvegicus clone CH230-272K4, *** SEQUENCING IN PROGRESS
DEFINITION *** 15 unordered pieces.
AC120454
AC120454.5 GI:25008669
VERSION HTG; HTGS PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
KEYWORDS Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus
ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Murioidea; Muridae; Murinae; Rattus;

1 (bases 1 to 246494)

Muzny,D.Marie., Metzker,M.Lee., Abramson,S., Adams,C., Alder,J.,

Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguilano,D.,

Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,

Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,

Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,

Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,

Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,

Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,

Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,

Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,

Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,

Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,

Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,

Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,

Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,

Gebregiorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,

Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,

Harvey,Y., Haylak,P., Hawes,A., Henderson,N., Hernanilton,K.,

Hernandez,R., Hines,S., Hladuk,S.L., Hodgson,A., Hogue,M.,

Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,

Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,

Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,

Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,

Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,

Lorensuhera,L., Louleghed,H., Lozado,R.J., Lu,X., Ma,J.,

Maheehwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,

Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,

Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,

Milosavljevic,A., Miner,G., Mirja,E., Montemayor,J., Moore,S.,

Morgan,M., Morris,K., Morris,S., Munidas,M., Murphy,M., Nair,L.,

Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,

Nwaokemele,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,

Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,

Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.,

Fuazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,

Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,

Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,

Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,

Shetty,J., Shvartabeyn,A., Sisson,I., Sitter,C.D., Smajls,D.,

Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,

Steinle,M., Strong,R., Sutton,A., Svatek,A., Taber,P., Taylor,C.,

Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,

Valas,R., Vera,V., Villalana,D., Waldron,L., Walker,B., Wang,J.,

Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,P.,

Williams,G., Willson,R., Wleczek,R., Wooden,H., Worley,K.,

Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,

Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von

Niederhauser,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,

Weinstock,G. and Gibbs,R.A.

Direct Submission

Unpublished

TITLE

JOURNAL

REFERENCE
AUTHORS
TITLE
JOURNAL

REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT

2 (bases 1 to 246494)
Worley,K.C.
Direct Submission
Submitted (07-MAY-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 246494)
Rat Genome Sequencing Consortium.
Direct Submission

Submitted (15-NOV-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Nov 15, 2002 this sequence version replaced gi:23611251.

The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: http://www.hgsc.bcm.tmc.edu/

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GWKZ

Center clone name: CH230-272K4

----- Summary Statistics

Assembly program: Phrap; version 0.990329

Consensus quality: 202515 bases at least Q40

Consensus quality: 207572 bases at least Q30

Consensus quality: 211369 bases at least Q20

Estimated insert size: 205757; sum-of-contigs estimation

Quality coverage: 5x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length

* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).

* NOTE: This is a 'working draft' sequence. It currently

* consists of 15 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

* as soon as it is available and the accession number will

* be preserved.

* 1 223610: contig of 223610 bp in length

* 223611 223710: gap of unknown length

* 223711 225139: contig of 1429 bp in length

* 225140 225239: gap of unknown length

* 225240 226743: contig of 1504 bp in length

* 226744 226843: gap of unknown length

* 226844 228228: contig of 1285 bp in length

* 228129 228228: gap of unknown length

* 228229 229753: contig of 1525 bp in length

* 229754 229853: gap of unknown length

* 229854 231650: contig of 1797 bp in length

* 231651 231750: gap of unknown length

* 231751 233063: contig of 1313 bp in length

* 233064 233163: gap of unknown length

* 233164 234274: contig of 1111 bp in length

* 234275 234374: gap of unknown length

* 234375 235660: contig of 1286 bp in length

* 235661 235760: gap of unknown length

* 235761 237660: contig of 1900 bp in length

* 237661 239922: gap of unknown length

* 239923 240022: contig of 2162 bp in length

* 240023 240022: gap of unknown length

* 240023 241516: contig of 1494 bp in length
 * 241517 241616: gap of unknown length
 * 241617 242741: contig of 1125 bp in length
 * 242742 242841: gap of unknown length
 * 242842 244151: contig of 1310 bp in length
 * 244152 244251: gap of unknown length
 * 244252 246494: contig of 2243 bp in length.

FEATURES

source 1..246494
 Location/Qualifiers
 misc_feature 1..246494
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-272K4"
 620..1244
 /note="clone boundary"
 /clone_end:Sp6
 site:
 end sequence: BZ240024"
 12216..13388
 /note="wgs contig"
 misc_feature complement(220115..220908)
 /note="clone boundary"
 /clone_end:T7
 site:
 end sequence: BZ240021"
 223611..223710
 /estimated_length=unknown
 gap 225140..225239
 /estimated_length=unknown
 gap 226744..226843
 /estimated_length=unknown
 gap 228129..228228
 /estimated_length=unknown
 gap 229754..229853
 /estimated_length=unknown
 gap 231651..231750
 /estimated_length=unknown
 gap 233064..233163
 /estimated_length=unknown
 gap 234275..234374
 /estimated_length=unknown
 gap 235661..235760
 /estimated_length=unknown
 gap 237661..237760
 /estimated_length=unknown
 gap 239923..240022
 /estimated_length=unknown
 gap 241517..241616
 /estimated_length=unknown
 gap 242742..242841
 /estimated_length=unknown
 gap 244152..244251
 /estimated_length=unknown

ORIGIN

Query Match 93.6%; Score 23.4; DB 14; Length 246494;
 Best Local Similarity 96.0%; Pred. No. 2.7;
 Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAAAACGACACGAGCCCTCG 25
 Db 231986 AAAAAACGACACGAGCCCTCG 232010

RESULT 5
 CP000079_06/c

WPCOMMENT

Sequence split into 12 fragments LOCUS CP000079 Accession CP000079
 Fragment Name Begin End
 CP000079_00 1 110000
 CP000079_01 100001 210000
 CP000079_02 200001 310000
 CP000079_03 300001 410000

CP000079_04 400001 510000
 CP000079_05 500001 610000
 CP000079_06 600001 710000
 CP000079_07 700001 810000
 CP000079_08 800001 910000
 CP000079_09 900001 1010000
 CP000079_10 1000001 1110000
 CP000079_11 1100001 1130447

Continuation (7 of 12) of CP000079 from base 600001 (CP000079 Leishmania major strain F

Query Match 76.8%; Score 19.2; DB 2; Length 110000;
 Best Local Similarity 87.5%; Pred. No. 2.9e+02;
 Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAACGACACGAGCCCTCG 24
 Db 14660 AAAAAACGACACGAGCCCTCG 14637

RESULT 6

AC113585

LOCUS

DEFINITION

AC113585

Trypanosoma brucei chromosome 5 clone RPI93-27M3, complete

sequence.

AC113585

AC113585.4 GI:62359100

HTG.

Trypanosoma brucei

Trypanosoma brucei

Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;

Trypanosoma.

REFERENCE

1 (bases 1 to 127380)

Authors

Ghedini, E., Blandin, G., Bartholomeu, D., Caler, E., Haas, B.,

Hannick, L., Shallom, J., Hou, L., Djikeng, A., Feldblyum, T.,

Hostetler, J., Johnson, J., Jones, K., Koo, H. L., Larkin, C.,

Peterson, J., Khalak, H. G., Salzberg, S., Simpson, A. J., Tallon, L.,

Van Aken, S., Wanless, D., White, O., Wortman, J., Fraser, C. M. and

El-Sayed, N. M. A.

Unpublished

2 (bases 1 to 127380)

Authors

El-Sayed, N. M., Khalak, H. and Adams, M. D.

Direct Submission

Submitted (02-MAR-2002) The Institute for Genomic Research, 9712

Medical Center Dr, Rockville, MD 20850, USA

JOURNAL

REFERENCE

3 (bases 1 to 127380)

Authors

El-Sayed, N. M., Khalak, H. and Adams, M. D.

Direct Submission

Submitted (05-APR-2002) The Institute for Genomic Research, 9712

Medical Center Dr, Rockville, MD 20850, USA

JOURNAL

REFERENCE

4 (bases 1 to 127380)

Authors

Haas, B., Blandin, G. and El-Sayed, N.

Direct Submission

Submitted (07-APR-2005) The Institute for Genomic Research, 9712

Medical Center Dr, Rockville, MD 20850, USA, nelsayed@tigr.org

On Apr 7, 2005 this sequence version replaced gi:19774505.

BAC clone RPI93-27M3 from Trypanosoma brucei chromosome 5.

Putative protein coding regions (>30 codons) were predicted by

combining the output of the gene finding algorithm GLIMMER with

similarity data and manual annotation of open reading frames.

Further refinement of gene models (additions, deletions and

alterations to boundary coordinates) were done manually based on

detailed comparisons with the T. cruzi and L. major genomes. Genes

on BACs were assigned a systematic name based on the chromosome and

the BAC from which they originated (e.g. Tb03.27F10.410). Gene

products were assigned a putative function when they shared

significant similarity with experimentally characterized gene

products or when they contained functionally known protein domains.

Gene products were labelled either 'hypothetical protein', or

'hypothetical protein, conserved' in the case of products showing

significant similarity with proteins of domains of unknown function

from other organisms. Gene products of unknown function with

predicted orthologs in Trypanosoma cruzi and Leishmania major

genomes were automatically classified as conserved proteins. Short

predicted coding sequences (less than 150 codons) with no additional evidence of coding potential were labelled 'hypothetical protein, unlikely'.

```
FEATURES
    source
        1..127380
            /organism="Trypanosoma brucei"
            /mol_type="genomic DNA"
            /isolate="GUTat10.1"
            /db_xref="taxon:5691"
            /chromosome="5"
            /clone="RPCI93-27M3"
        1..672
            /note="annotation for this sequence region can be found on
            overlapping clone 26K5."
        <1078..>2694
            /locus_tag="Tb927.5.3070"
            /note="synonym: Tb05.27M3.20; hypothetical protein,
            conserved"
        <1078..>2694
            /locus_tag="Tb927.5.3070"
            /product="hypothetical protein, conserved"
            1078..2694
            /locus_tag="Tb927.5.3070"
            /note="go function: zinc ion binding [goid 0008270]"
            /codon_start=1
            /product="hypothetical protein, conserved"
            /protein_id="AAK79547.1"
            /db_xref="GI:62359101"
            /translation="MAKKKSRGVEVKDISNFOYGYRADSSNAVTPVARRVPVDP
            AVLKRAYILHNFQVILRHDAQOLRAVPGLPORPHOVTEALRLSGTTPPWDLIIAIVAR
            GTEEPHCPCLEEQAARITSCGVFCLVCVLOYISRRKASKORTCPVCNFIIVA
            SLRCPMVRVLPSPVGARASFTMLKRGKSCILLQDDPCWKETLTDDDELRLPFVGE
            PSATYSYVISTEGATEUREDECTAIRTKQISMLEAQRPFTSFDDLLKRAEDAIRL
            TLQGVPPQASGETSPPLAPKTGTETNMAYEFYGTGQAYIYHPIITYKMLCVDA
            EARNSPLRVVAPVEEITFTQDEASRKYRVFAHPLHATIKLCLDLSDIVLAST
            MKIFETLARLKRQLRESSISSTEDSTWOEYLRRVRTDWGAEARTLEDMSPMSE
            FMWSEADFERLPALELSPSGTPLGPVQASAGRETPGGCGWKTNLSLQLFSSPP
            VAPVPTWGGHGFENKPNSSNVSLEGRKR"
        2852..2879
            /note="microsatellite"
            /rpt_type=tandem
            /rpt_unit="ta"
            <3395..>4403
            /locus_tag="Tb927.5.3080"
            /note="synonym: Tb05.27M3.40; hypothetical protein,
            conserved"
            <3399..>4403
            /locus_tag="Tb927.5.3080"
            /product="hypothetical protein, conserved"
            3399..4403
            /locus_tag="Tb927.5.3080"
            /codon_start=1
            /product="hypothetical protein, conserved"
            /protein_id="AAK79548.1"
            /db_xref="GI:62359102"
            /translation="MWGPQEVANPHRDKGSFVLPSSTAAGAAALCASLLPHSNLVCA
            PLNPHIAREPDTATGGDGTQVALLTRFRIENERLDDQVKAHVQVQALLTLRLQELS
            ARNTALHQDFRRKNDNAVVEKLKKQLEVMEVRAALQVGLNGRLSAYBEAISRPH
            DHPQGLSANMDRVRELI CEALSAGFLTRVANAVEHCHPSTHEVTDCCPSVVEECTNH
            CQAAMRGKELPGETAARGHHVSCCELASALREEVLFCTEAVRVAAKLPWERSAD
            SSNAADRGTGLQHESSAPLARPTAECAEAHELQPSDRRPRRPAARPTVEDCEVQ"
        <4874..>5179
            /locus_tag="Tb927.5.3090"
            /note="synonym: Tb05.27M3.50; hypothetical protein,
            conserved"
            <4874..>5179
            /locus_tag="Tb927.5.3090"
            /product="hypothetical protein, conserved"
            4874..5179
            /locus_tag="Tb927.5.3090"
            /codon_start=1
            /product="hypothetical protein, conserved"
```

```
repeat_region
    repeat_region
        8385..8409
            /protein_id="AAK79549.1"
            /db_xref="GI:62359103"
            /translation="WSSGFHPDVSNDAIKMPPEALHKKLENAQLAHRICLAKALK
            AGEPPVKCALTWGEYLIRYQAWSEYRFPFQDSVAQAKYKYSWKKRQEEDDKNFK"
        8385..8409
            /note="microsatellite"
            /rpt_type=tandem
            /rpt_unit="tg"
        9246..9265
            /note="microsatellite"
            /rpt_type=tandem
            /rpt_unit="ta"
            complement(<9330..>10094)
            /locus_tag="Tb927.5.3100"
            /note="synonym: Tb05.27M3.80; hypothetical protein"
            complement(<9330..>10094)
            /locus_tag="Tb927.5.3100"
            /product="hypothetical protein"
            complement(9330..10094)
            /locus_tag="Tb927.5.3100"
            /codon_start=1
            /product="hypothetical protein"
            /protein_id="AAK79550.1"
            /db_xref="GI:62359104"
            /translation="MINKSGRVHYSTDSSVSGSKQITHGGIGKSGRYNINPAIAW
            LDETMLKDDPKLYTNHSTVEBOQGRMWSLETETTYQFITLLQEYYPWVGRNVGADE
            VRRTIIDLREALSRDRTTIKDLEAGSOLFRLIFGYENDWIRLELGSAGOLKRRHLHT
            EKMIYVGLKLSALYIMNVNKGKQKNSCDGILRLLEELSRSSGVNMLPPANPENSFP
            PPSVNGHPQTLPRALRNDELQAFATVTPDFEIRGE"
        10667..10686
            /note="microsatellite"
            /rpt_type=tandem
            /rpt_unit="gaaaa"
            complement(<10702..>11358)
            /locus_tag="Tb927.5.3110"
            /note="synonym: Tb05.27M3.100; hypothetical protein,
            conserved"
            complement(<10702..>11358)
            /locus_tag="Tb927.5.3110"
            /product="hypothetical protein, conserved"
            complement(10702..11358)
            /locus_tag="Tb927.5.3110"
            /codon_start=1
            /product="hypothetical protein, conserved"
            /protein_id="AAK79551.1"
            /db_xref="GI:62359105"
            /translation="MLNRSRLRRVWLGTAANTSGLRDVMPSTGKYNIVSPTDVAAN
            ISAGAMPEANVAALPVRIQYSKNRALRSDVYDELLKPLRYKLHDFAKLCPPQR
            COEPEAVGVVGRDSAVGYLPLGPADPLDTIPFFVHRSSNGVLPGKFSMHRNIMP
            AFYLRIQQVEGDMRFEPELLKIPTTKKIFVRSHSIYVYNVALDGRMVLHWWLLGLGF"
        11620..11647
            /note="microsatellite"
            /rpt_type=tandem
            /rpt_unit="ta"
            12205..12252
            /note="microsatellite"
            /rpt_type=tandem
            /rpt_unit="ta"
            complement(<12339..>13280)
            /locus_tag="Tb927.5.3120"
            /note="synonym: Tb05.27M3.120; translation initiation
            factor, putative"
            complement(<12339..>13280)
            /locus_tag="Tb927.5.3120"
            /product="translation initiation factor, putative"
            complement(12339..13280)
            /locus_tag="Tb927.5.3120"
            /note="go component: cytoplasm [goid 00057371];
            go_function: translation initiation factor activity [goid
            00037431];
            go_process: translational initiation [goid 0006413]"
            /codon_start=1
```

```
/product="translation initiation factor, putative"
/protein_id="AAK79552.1"

Query Match      76.8%;   Score 19.2;   DB 2;   Length 127380;
Best Local Similarity 87.5%;   Pred. No. 2.9e+02;
Matches 21;   Conservative 0;   Mismatches 3;   Indels 0;   Gaps 0;

QY 1 AAAAAACGACACGAGCGCCCTCG 24
    |||||
Db 94165 AAAAAACGACACGCGAACCCCTCG 94188
    |||||

RESULT 7
BC071044      2426 bp      mRNA      linear      VRT 19-JUL-2004
LOCUS
DEFINITION
Xenopus laevis hypothetical protein MGC83180, mRNA (cDNA clone
MGC:83180 IMAGE:6631780), complete cds.
ACCESSION
BC071044
VERSION
BC071044.1 GI:47506955
KEYWORDS
MGC.
SOURCE
Xenopus laevis (African clawed frog)
ORGANISM
Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
Xenopodinae; Xenopus; Xenopus.
1 (bases 1 to 2426)
Klein, S.L., Strausberg, R.L., Wagner, L., Pontius, J., Clifton, S.W.
and Richardson, P.
Genetic and genomic tools for Xenopus research: The NIH Xenopus
Initiative
Dev. Dyn. 225 (4), 384-391 (2002)
12454917
2 (bases 1 to 2426)
Klausner, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Schlauser, R.D., Collins, F.S., Wagner, L., Shennan, C.M., Schuler, G.D.,
Altshuler, S.P., Zebrow, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Scapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
Abramson, R.D., Prange, C., Raha, S., Loquellano, N.A., Peters, G.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettman, M., Madan, A., Rodrigues, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalios, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
3 (bases 1 to 2426)
Klein, S. and Strausberg, R.
Direct Submission
Submitted (18-MAY-2004) National Institutes of Health, Xenopus Gene
Collection (XGC), National Institute of Child Health and Human
Development, 6100 Executive Boulevard, Room 4B01, Rockville, MD
20892-7510, USA
NIH-XGC Project
Contact: XGC help desk
Email: cgapbe-remail.nih.gov
Tissue Procurement: Martha Rebbert, Steven L. Klein, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILLNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca
Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Cloutier, Ruth
Andy Chan, Steve S. Chand, William Chow, Alison Butterfield,
```

Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao, Kim MacDonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Morin, Teika Olson, Diana Palmquist, Anca Petrescu, Anna Lilia Prabhau, Parvaneh Saeedi, JR Santos, Angeliue Schnerch, Ursula Skalska, Duane Smalios, Jeff Stott, Miranda Teai, George Yang, Jacquie Schein, Asim Siddiqui, Rob Holt, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Image: 157 Row: a Column: 21
This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein.

```
FEATURES
Location/Qualifiers
1..2426
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone="MGC:83180 IMAGE:6631780"
/tissue_type="Embryo, stage 10"
/clone_lib="NICHG_XGC_Emb1"
/lab_host="DH10B"
/notes="Vector: PCMV-SPORT6"
1..2426
/gene="MGC83180"
/db_xref="GeneID:432098"
89..946
/gene="MGC83180"
/codon_start=1
/product="MGC83180 protein"
/protein_id="AAH71044.1"
/db_xref="GI:47506955"
/db_xref="GeneID:432098"
/translation="MREVLQOQHPGTVEPPGLFRLEAHRIVISLGIYHSRVQ
RGGLKLNLMVSLVLSAQVYLSQSPLEQYVLRQALHNPQDCKEPCPL
TEGTCPPQEQDARTCTSCPTASRGNPDNTENPRGVCQGHVTSRSHSLHEP
DICPSHEAPSPSPSSCCRRKGNPDGPMKRAKREQSGQVEPFGDEAEHME
TENVANLISIFGTSFSLSKPEPTAEAEPAQDVQELDSTGAPGCSERAISSMS
SNSTAIEAF"
```

ORIGIN

```
Query Match      75.2%;   Score 18.8;   DB 5;   Length 2426;
Best Local Similarity 90.9%;   Pred. No. 5.5e+02;
Matches 20;   Conservative 0;   Mismatches 2;   Indels 0;   Gaps 0;

QY 4 AAACGACACGAGCGCCCTCG 25
    |||||
Db 478 AAACGACACAGAACCCCTCG 499
    |||||
```

RESULT 8

```
AC155827/c
LOCUS
DEFINITION
Mus musculus 10 BAC RP24-408N5 (Roswell Park Cancer Institute
(C57BL/6J Male) Mouse BAC Library) complete sequence.
```

```
ACCESSION
AC155827 AC116864
VERSION
AC155827.10 GI:67972468
KEYWORDS
HTG.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
```

REFERENCE

```
AUTHORS
1 (bases 1 to 149754)
Muzny, D., Adams, C., Agbai II, O., Allen, C., Alsbrooks, S., Archer, P.,
Arredondo, H., Bandaranaike, D., Bangura, L., Beltran, B., Beltran, R.,
Beraducci, A., Biswalo, K., Blyth, P., Bonham, H., Buhay, C., Burch, P.,
Cadoree, I., Canada, A., Cardenas, V., Carter, K., Cavazos, I.,
Chacko, J., Chahrour, M., Chavez, D., Chen, A., Chen, G., Chen, R.,
Cheng, M.-T., Chu, J., Clerc, K., Cockrell, R., Coyle, M., Cree, A.,
Curry, S., Dai, W., Davila, M.L., Davis, C., Davy-Carroll, L., De
Anda, C., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H.,
Donlin, J., McCauley, S., Dugan-Rocha, S., Dunn, A., Durbin, K.,
```

Dziuda, D., Egan, A., Escotto, M., Espinosa, V., Eugene, C., Fa, M., Fernandez, S., Fernando, P., Flagg, N., Forbes, L., Foster, P., Fowler, G., Fu, Q., Fuh, E., Garcia, A., Garcia, R., Garner, T., Gaskin, C., Gench, S., Ghose, S., Gill, R., Gonzalez, D., Gonzalez-Garay, M., Guevara, W., Holder, M., Haaland, W., Haerberlein, K., Hall, B., Hamid, H., Hamilton, K., Harbes, B., Harris, R., Havlak, P., Hawes, A., Hawkins, E., Hayes, S., Hemphill, L., Hernandez, J., Hines, S., Hitchens, M., Hodgson, A., Hogues, M., Hollins, B., Howell, L. T., Hulyk, S., Hume, J., Ima, K., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Kalafus, K., Kelly, S., Keys, T., Khan, Z., King, L., Kovar, C., Kowis, A., Kowis, C., Lara, F., Leal, S., Lee, K., Lee, S., Legall, F. I., Lemon, S., Lewis, L., Li, B., Li, Y., Li, Z., Linell, M., Liu, W., Liu, Y. S., Liu, Y., Liyanage, D., London, P., Lopez, J., Lorenshew, L., Lozado, R., Luk, T., Madu, R., Maheshwari, M., Mahoney, C., Malloy, K., Mansouri, D., Martinez, E., McClelland, H., McPherson, J., Mercadao, C., Metzger, M., Milosavljevic, A., Minja, E., Morgan, M., Morris, S., Munidasa, M., Murray, D., Nazarith, L., Ngo, D., Nguyen, N., Norwig-Eastaugh, E., Nott, A., Nwaokemele, O., Obregon, M., Ochi-Okorie, C., Odeh, E., Okwuonu, G., Okwuonu, K., Parker, D., Pasternak, S., Patel, B., Patel, V., Paul, H., Perez, A., Perez, L., Petrosino, J., Pham, T., Primus, E., Pu, L., Puazo, M., Qin, X., Quinn, A., Quiroz, J., Rabata, D., Rachlin, E., Reigh, R., Ren, Y., Reuter, M., Richards, S., Rives, C., Rodriguez, F., Rojas, A., Ruiz, S. J., Sana, M., Sanders, W., Santibanez, J., Santos, R., Saverly, G., Scherer, S., Shen, H., Shen, Y., Sisson, I., Sneed, A., Sodergren, E., Song, X. Z., Sorelle, R., Svatek, A., Taylor, E., Taylor, T., Thomas, N., Thorn, R., Thornton, R., Trejos, Z., Usmani, K., Vargo, C., Verduzco, D., Villaseña, D., Virk, D., Volkov, A., Waldron, L., Walker, B., Wang, Q., Wang, S., Warren, J., Wei, X., Wheeler, D., Williams, G., Williams, R., Worley, K., Wright, R., Wu, J., Yakub, S., Yan, K., Yuan, Y., Yu, F., Zhang, J., Zhang, L., Zhang, Z., Zhou, J., Weinstock, G. and Gibbs, R.

Direct Submission
Unpublished
2 (bases 1 to 149754)
Worley, K.C.
Direct Submission
Submitted (20-JAN-2005) Human Genome Sequencing Center, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 149754)
Worley, K.C.
Direct Submission
Submitted (01-MAR-2005) Human Genome Sequencing Center, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
4 (bases 1 to 149754)
Worley, K.C.
Direct Submission
Submitted (11-MAR-2005) Human Genome Sequencing Center, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
5 (bases 1 to 149754)
Worley, K.C.
Direct Submission
Submitted (18-JUN-2005) Human Genome Sequencing Center, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Jun 18, 2005 this sequence version replaced gi:60735107.
Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality does not meet this standard, it will be indicated in the annotation.

The repeat regions shown were identified using RepeatMasker by Adrian Smit.

Sequence similarities were identified using Powerblast by Jinghui Zhang.

Exon/Intron boundaries of identified genes were chosen if there were canonical splice junctions that maintained sequence continuity across the splice junctions.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM

repeat_region	/rpt family="AT rich" complement(12248..12353) /rpt family="PB1D10" 13905..14043	JOURNAL REFERENCE AUTHORS TITLE JOURNAL COMMENT	Submitted (29-JAN-2004) NIH Intramural Sequencing Center, 8717 Groveomnt Circle, Gaithersburg, MD 20877, USA 3 (bases 1 to 163916) Green, E.D. Direct Submission Submitted (15-APR-2004) NIH Intramural Sequencing Center, 8717 Groveomnt Circle, Gaithersburg, MD 20877, USA On Apr 15, 2004 this sequence version replaced gi:41386836. ----- Genome Center Center: NIH Intramural Sequencing Center Center code: NISC Web site: http://www.nisc.nih.gov Contact: nisc.zoo@nhgri.nih.gov ----- Project Information Center project name: flyg Center clone name: 030P02
repeat_region	/rpt family="B2_Mm2" 14951..15017		
repeat_region	/rpt family="TA)n" 15782..15814		
repeat_region	/rpt family="AT rich" complement(16006..16145) /rpt family="Blf" 17478..17501		
repeat_region	/rpt family="AT rich" 18588..18644		
repeat_region	/rpt family="(TCCA)n" complement(18653..18795) 19137..19223		
repeat_region	/rpt family="RSINE1" 19348..19406		
repeat_region	/rpt family="(TG)n" 19576..19725		
repeat_region	/rpt family="Bl_Mur3" 19726..19762		
repeat_region	/rpt family="GA-rich" 20018..20083		
repeat_region	/rpt family="PB1D10" complement(20257..20326) /rpt family="ID4" 20793..20819		
repeat_region	/rpt family="(CAAAA)n" 21389..21428		
repeat_region	/rpt family="A-rich"		
Query Match	75.2%; Score 18.8; DB 9; Length 149754;		
Best Local Similarity	90.9%; Pred. No. 4.5e+02;		
Matches	20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;		
Qy	2 AAAAAACACACAGCGAGCCCTC 23 		
Db	80249 AAAAAACACACAGCGAGCCCTC 80228		
RESULT 9			
AC148067			
LOCUS	AC148067 163916 bp DNA linear HTG 15-APR-2004		
DEFINITION	Canis familiaris clone RP81-30P2, WORKING DRAFT SEQUENCE, 9 ordered pieces.		
ACCESSION	AC148067		
VERSION	AC148067.2 GI:46391174		
KEYWORDS	HTG; HTGS_PHASE2; HTGS_DRAFT.		
SOURCE	Canis familiaris (dog)		
ORGANISM	Canis familiaris Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae; Canis.		
REFERENCE	1 (bases 1 to 163916) Antonellis, A., Ayale, K., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S., Chu, G., Coleman, B., Coleman, H., Daki, N., Engle, J., Granite, S., Guan, X., Gupta, J., Haghighi, P., Han, J., Hansen, N., Ho, S.-L., Hu, P., Hurle, B., Idol, J.R., Jones, C., Karlins, E., Kim, H., Kwong, P., Laric, P., Larson, S., Lee-Lin, S.-O., Legaspi, R., Maduro, Q.L., Maduro, V.B., Margulies, E.H., Masiello, C., Maskeri, B., McDowell, J., Mullikin, J.C., Pagnier, C., Portnoy, M.E., Prasad, A., Puri, O., Reddi-Dugue, N., Schandler, K., Schueler, M.G., Shah, K., Sison, C., Standripop, S., Thomas, J.W., Thomas, P.J., Tsipouri, V., Vogt, J.L., Wetherby, K.D., Young, A. and Green, E.D.		
AUTHORS	NISC Comparative Sequencing Initiative Unpublished 2 (bases 1 to 163916) Green, E.D. Direct Submission	FEATURES source	Location/Qualifiers 1..163916 /organism="Canis familiaris" /mol_type="genomic DNA" /db_xref="taxon:9615" /clone="RP81-30P2"

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.950319
Consensus quality: 162694 bases at least Q40
Consensus quality: 162942 bases at least Q30
Consensus quality: 163053 bases at least Q20
Insert size: 149000; agarose-fp
Insert size: 163116; sum-of-contigs
Quality coverage: 11.40x in Q20 bases; agarose-fp
Quality coverage: 10.42x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently consists of 9 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.

* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 10164: contig of 10164 bp in length
* 10165 10264: gap of unknown length
* 10265 13805: contig of 3541 bp in length
* 13806 13905: gap of unknown length
* 13906 16621: contig of 2716 bp in length
* 16622 16721: gap of unknown length
* 16722 60680: contig of 43959 bp in length
* 60681 60780: gap of unknown length
* 60781 78425: contig of 17645 bp in length
* 78426 78526: gap of unknown length
* 78527 100161: contig of 21636 bp in length
* 100162 100261: gap of unknown length
* 100262 103084: contig of 2823 bp in length
* 103085 103184: gap of unknown length
* 103185 105651: contig of 2467 bp in length
* 105652 105751: gap of unknown length
* 105752 163916: contig of 58165 bp in length.

```
/clone_lib="RP81"
/note="BAC resource: http://bacpac.chori.org/"
1..10164
/note="assembly_fragment
clone_end:SP6
vector_side:left"
10165..10264
/estimated_length=unknown
10265..13805
/note="assembly_fragment"
13806..13905
/estimated_length=unknown
13906..16621
/note="assembly_fragment"
16622..16721
/estimated_length=unknown
16722..60680
/note="assembly_fragment"
60681..60780
/estimated_length=unknown
60781..78425
/note="assembly_fragment"
78426..78525
/estimated_length=unknown
78526..100161
/note="assembly_fragment"
100162..10261
/estimated_length=unknown
10262..103084
/note="assembly_fragment"
103085..103184
/estimated_length=unknown
103185..105651
/note="assembly_fragment"
105652..105751
/estimated_length=unknown
105752..163916
/note="assembly_fragment
clone_end:r7
vector_side:right"

ORIGIN
Query Match 75.2%; Score 18.8; DB 14; Length 163916;
Best Local Similarity 90.9%; Pred. No. 4.5e+02;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAACGACACAGCGAGCCCT 22
||||| ||||| ||||| |||||
Db 21623 AAAAAACGACACAGAGCCCT 21644

RESULT 10
CR386415/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Gallus gallus (chicken)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 791)
REFERENCE
AUTHORS
Boardman,P.B., Bonfield,J.K., Brown,W.R.A., Carder,C., Chalk,S.E.,
Croning,M.D.R., Davies,R.M., Francis,M.D., Grafham,D.V.,
Hubbard,S.J., Humphray,S.J., Hunt,P.J., Maddison,M., McLaren,S.R.,
Niblett,D., Overton,I.M., Rogers,J., Scott,C.E., Taylor,R.G.,
Tickle,C. and Wilson,S.A.
Direct Submission
Submitted (05-APR-2004) Sanger Institute, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: chickest@bms.umist.ac.uk
BBSRC/Dundee/Nottingham/Sanger/Sheffield/UMIST Gallus gallus cDNA

sequencing project.
This sequence is from the
BBSRC/Dundee/Nottingham/Sanger/Sheffield/UMIST cDNA collection,
from a library constructed by Elizabeth Bosch. cDNA was prepared
from RNA extracted from chondrocytes isolated from growth plates,
normalised,
and poly A-trimmed. EcoRI-NotI cut cDNA was then ligated into the
vector. Vector: pBluescript II KS(+); Site_1: EcoRI; Site_2: NotI
Host: Escherichia coli DH10B.
Location/Qualifiers
1..791
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="Layer and broiler"
/db_xref="taxon:9031"
/clone="ChEST231n18"
/clone_lib="CSEQRBN09"
/dev_stage="juvenile"

FEATURES
source
1..791
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="Layer and broiler"
/db_xref="taxon:9031"
/clone="ChEST231n18"
/clone_lib="CSEQRBN09"
/dev_stage="juvenile"

ORIGIN
Query Match 74.4%; Score 18.6; DB 5; Length 791;
Best Local Similarity 84.0%; Pred. No. 7.3e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAACGACACAGCGAGCCCTCGG 25
||||| ||||| ||||| |||||
Db 736 AAAAAATGACAAAGCGAGCCTTCG 702

RESULT 11
AE013307
LOCUS
DEFINITION
Methanosarcina mazei strain Goel, section 89 of the complete
genome.
ACCESSION
AE013307 AE008384
VERSION
AE013307.1 GI:20905228
KEYWORDS
SOURCE
Methanosarcina mazei Goel
ORGANISM
Methanosarcina mazei Goel
Archaea; Euryarchaeota; Methanomicrobia; Methanosarcinales;
Methanosarcinaceae; Methanosarcina.
1 (bases 1 to 10029)
REFERENCE
AUTHORS
Deppenmeier,U., Johann,A., Hartsch,T., Merkl,R., Schmitz,R.A.,
Martinez-Arias,R., Henne,A., Wierer,A., Baeumer,S., Jacobi,C.,
Brueggemann,H., Lienard,T., Christmann,A., Boencke,M., Steckel,S.,
Bhattacharyya,A., Lykidis,A., Overbeek,R., Klenk,H.-P.,
Gunsalus,R.P., Fritz,H.-J. and Gottschalk,G.
The genome of Methanosarcina mazei: evidence for lateral gene
transfer between bacteria and archaea
J. Mol. Microbiol. Biotechnol. 4 (4), 453-461 (2002)
12125824
2 (bases 1 to 10029)
Deppenmeier,U., Johann,A., Hartsch,T., Merkl,R., Schmitz,R.A.,
Martinez-Arias,R., Henne,A., Wierer,A., Baeumer,S., Jacobi,C.,
Brueggemann,H., Lienard,T., Christmann,A., Boencke,M., Steckel,S.,
Bhattacharyya,A., Lykidis,A., Overbeek,R., Klenk,H.-P.,
Gunsalus,R.P., Fritz,H.-J. and Gottschalk,G.
Direct Submission
Submitted (23-AUG-2001) Goettingen Genomics Laboratory, Institute
for Microbiology and Genetics, Grisebachstrasse 8, Goettingen
37077, Germany
Location/Qualifiers
1..10029
/organism="Methanosarcina mazei Goel"
/mol_type="Genomic DNA"
/strain="Goel"
/db_xref="taxon:192952"
complement (121..1278)
/gene="NM0818"
complement (121..1278)
/gene="NM0818"
complement (121..1278)
/gene="NM0818"
/EC_number="2.7.2.11"
/codon_start=1

FEATURES
source
1..10029
/organism="Methanosarcina mazei Goel"
/mol_type="Genomic DNA"
/strain="Goel"
/db_xref="taxon:192952"
complement (121..1278)
/gene="NM0818"
complement (121..1278)
/gene="NM0818"
/EC_number="2.7.2.11"
/codon_start=1

gene
CDS
```


RESULT 12	RESULT 13
AC100392	AC100392
LOCUS	LOCUS
DEFINITION	DEFINITION
ACCESSION	ACCESSION
VERSION	VERSION
KEYWORDS	KEYWORDS
SOURCE	SOURCE
ORGANISM	ORGANISM
REFERENCE	REFERENCE
AUTHORS	AUTHORS
TITLE	TITLE
JOURNAL	JOURNAL
COMMENT	AUTHORS

<p> BX571685 Mouse DNA sequence from clone Rp23-29D24 on chromosome 4, complete sequence. BX571685 BX571685.2 GI:46917204 HTG. Mus musculus (house mouse) Mus musculus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidae; Muridae; Murinae; Mus. 1 (bases 1 to 54569) Botcherby, M. Direct Submission Submitted (30-APR-2004) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk On Apr 30, 2004 this sequence version replaced gi:32567440. </p>	<p> BX571685 Mouse DNA sequence from clone Rp23-29D24 on chromosome 4, complete sequence. BX571685 BX571685.2 GI:46917204 HTG. Mus musculus (house mouse) Mus musculus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidae; Muridae; Murinae; Mus. 1 (bases 1 to 54569) Botcherby, M. Direct Submission Submitted (30-APR-2004) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk On Apr 30, 2004 this sequence version replaced gi:32567440. </p>
---	---

```

ORIGIN
Query Match          74.4%; Score 18.6; DB 9; Length 54589;
Best Local Similarity 84.0%; Pred. No. 5.9e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      1  AAAAAACGACACAGCGCCCTCGG 25
         ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db       1048 AAAAAATGACCCAGCTAGCCCTCTG 1072

```

AC100392 65758 bp DNA linear HTG 22-NOV-2001
Mus musculus clone RP23-133K10, LOW-PASS SEQUENCE SAMPLING.
AC100392 AC100392.1 GI:17047758
HTG; HTGS PHASE0.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidae; Muridae; Murinae; Mus.
1 (bases 1 to 65758)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Mus musculus, clone RP23-133K10
Unpublished
2 (bases 1 to 65758)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Boguslavskiy,L., Boukhgalter,B.,
Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
Colangelo,K., Collins,S., Collymore,A., Cook,A.,
Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliiev,I., Johnson,R.,
Jones,C., Katat,A., Karatas,A., Kells,C., LaRoque,K.,
Llamazares,R., Landers,T., Lehoczký,J., Levine,R., Liu,G.,
MacLean,C., MacDonald,P., Major,J., Marguis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J.,
Meneus,L., Mihova,T., Menga,V., Murphy,T., Naylor,J., Nguyen,C.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupbach,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Trophank,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Vainel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zinkel,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (22-NOV-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:
[Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: submissions@genome.wi.mit.edu
----- Project Information
Center project name: L15260
Center clone name: 133_K_10

* NOTE: This record contains 81 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
* 1 709: contig of 709 bp in length
* 710 809: gap of 100 bp
* 810 1526: contig of 717 bp in length
* 1527 1626: gap of 100 bp
* 1627 2353: contig of 727 bp in length
* 2354 2454: gap of 100 bp
* 2454 3259: contig of 806 bp in length
* 3260 3359: gap of 100 bp

```
* 3360 4096: contig of 737 bp in length
* 4097 gap of 100 bp
* 4197 contig of 721 bp in length
* 4918 gap of 100 bp
* 5018 contig of 697 bp in length
* 5018 contig of 697 bp in length
* 5018 gap of 100 bp
* 5815 contig of 743 bp in length
* 5857 gap of 100 bp
* 6558 contig of 687 bp in length
* 7344 gap of 100 bp
* 7345 gap of 100 bp
* 7445 contig of 726 bp in length
* 8171 gap of 100 bp
* 8271 contig of 726 bp in length
* 8996 gap of 100 bp
* 8997 gap of 100 bp
* 9097 contig of 706 bp in length
* 9803 gap of 100 bp
* 9903 contig of 731 bp in length
* 10634 gap of 100 bp
* 10734 contig of 683 bp in length
* 11417 gap of 100 bp
* 11517 contig of 712 bp in length
* 12229 gap of 100 bp
* 12329 contig of 711 bp in length
* 13040 gap of 100 bp
* 13139 contig of 721 bp in length
* 13140 contig of 721 bp in length
* 13861 gap of 100 bp
* 14674 contig of 714 bp in length
* 14675 gap of 100 bp
* 14775 contig of 715 bp in length
* 15489 gap of 100 bp
* 15589 contig of 731 bp in length
* 15590 contig of 731 bp in length
* 16320 gap of 100 bp
* 16421 contig of 686 bp in length
* 17107 gap of 100 bp
* 17207 contig of 697 bp in length
* 17904 gap of 100 bp
* 18003 contig of 737 bp in length
* 18004 gap of 100 bp
* 18741 contig of 707 bp in length
* 18841 gap of 100 bp
* 19548 gap of 100 bp
* 19648 contig of 661 bp in length
* 20308 gap of 100 bp
* 20409 contig of 716 bp in length
* 21124 gap of 100 bp
* 21125 contig of 715 bp in length
* 21225 gap of 100 bp
* 21940 contig of 660 bp in length
* 22040 gap of 100 bp
* 22700 contig of 753 bp in length
* 22800 gap of 100 bp
* 23553 contig of 719 bp in length
* 24371 gap of 100 bp
* 24372 contig of 702 bp in length
* 24472 gap of 100 bp
* 25174 gap of 100 bp
* 25274 contig of 701 bp in length
* 25975 gap of 100 bp
* 26075 contig of 715 bp in length
* 26889 gap of 100 bp
* 26890 contig of 722 bp in length
* 27612 gap of 100 bp
* 27712 contig of 716 bp in length
* 28428 gap of 100 bp
* 28528 contig of 724 bp in length
* 29252 gap of 100 bp
* 29352 contig of 702 bp in length
* 30053 gap of 100 bp
* 30153 contig of 696 bp in length
* 30154 gap of 100 bp
* 30949 gap of 100 bp
* 30950 contig of 587 bp in length
* 31537 gap of 100 bp
* 31637 contig of 733 bp in length
* 32370 gap of 100 bp
* 32470 contig of 719 bp in length
```

```
* 33189 33288: gap of 100 bp
* 33289 contig of 719 bp in length
* 34008 gap of 100 bp
* 34107 contig of 740 bp in length
* 34108 gap of 100 bp
* 34848 contig of 735 bp in length
* 34948 gap of 100 bp
* 35682 contig of 710 bp in length
* 35783 gap of 100 bp
* 36492 contig of 726 bp in length
* 36493 gap of 100 bp
* 37318 contig of 706 bp in length
* 37319 gap of 100 bp
* 37419 contig of 706 bp in length
* 38125 gap of 100 bp
* 38224 contig of 695 bp in length
* 38919 gap of 100 bp
* 39019 contig of 708 bp in length
* 39727 gap of 100 bp
* 39827 gap of 100 bp
* 39828 contig of 638 bp in length
* 40466 gap of 100 bp
* 40566 contig of 748 bp in length
* 41131 gap of 100 bp
* 41134 gap of 100 bp
* 41414 contig of 720 bp in length
* 42133 gap of 100 bp
* 42233 contig of 728 bp in length
* 42961 gap of 100 bp
* 43061 contig of 700 bp in length
* 43062 gap of 100 bp
* 43762 contig of 710 bp in length
* 43862 gap of 100 bp
* 44571 gap of 100 bp
* 44572 contig of 721 bp in length
* 45392 gap of 100 bp
* 45492 contig of 699 bp in length
* 46191 gap of 100 bp
* 46192 contig of 689 bp in length
* 46292 gap of 100 bp
* 46980 contig of 622 bp in length
* 47080 gap of 100 bp
* 47702 contig of 681 bp in length
* 47802 gap of 100 bp
* 47803 contig of 694 bp in length
* 48484 gap of 100 bp
* 48584 contig of 694 bp in length
* 49277 gap of 100 bp
* 49377 contig of 708 bp in length
* 50085 gap of 100 bp
* 50186 contig of 688 bp in length
* 50873 gap of 100 bp
* 50974 contig of 728 bp in length
* 50975 gap of 100 bp
* 51701 contig of 729 bp in length
* 51801 gap of 100 bp
* 52530 contig of 707 bp in length
* 52630 gap of 100 bp
* 52631 contig of 703 bp in length
* 53337 gap of 100 bp
* 53338 contig of 703 bp in length
* 53438 gap of 100 bp
* 54141 contig of 709 bp in length
* 54241 gap of 100 bp
* 54950 contig of 718 bp in length
* 55050 gap of 100 bp
```

Query Match 74.4%; Score 18.6; DB 14; Length 65758;
Best Local Similarity 84.0%; Pred. No. 5.8e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 AAAAAACACACACGAGCCCTCGG 25
Db 58799 AAAAAACGATCATCGGCGCCCTCGG 58823

RESULT 14

AL845429/c
LOCUS AL845429
DEFINITION Mouse DNA sequence from clone RP23-29004 on chromosome 2, complete sequence.
ACCESSION AL845429
VERSION AL845429.4 GI:23895741

97626 bp DNA linear ROD 09-OCT-2002
AL845429
Mouse DNA sequence from clone RP23-29004 on chromosome 2, complete sequence.
AL845429
AL845429.4 GI:23895741

RESULT 17			
AP008216_225/c			
WPCOMMENT			
Sequence split into 227 fragments			
LOCUS AP008216 Accession AP008216			
Fragment Name	Begin	End	
AP008216_000	1	110000	
AP008216_001	100001	210000	6800001
AP008216_002	200001	310000	6900001
AP008216_003	300001	410000	7000001
AP008216_004	400001	510000	7100001
AP008216_005	500001	610000	7200001
AP008216_006	600001	710000	7300001
AP008216_007	700001	810000	7400001
AP008216_008	800001	910000	7500001
AP008216_009	900001	1010000	7600001
AP008216_010	1000001	1110000	7700001
AP008216_011	1100001	1210000	7800001
AP008216_012	1200001	1310000	7900001
AP008216_013	1300001	1410000	8000001
AP008216_014	1400001	1510000	8100001
AP008216_015	1500001	1610000	8200001
AP008216_016	1600001	1710000	8300001
AP008216_017	1700001	1810000	8400001
AP008216_018	1800001	1910000	8500001
AP008216_019	1900001	2010000	8600001
AP008216_020	2000001	2110000	8700001
AP008216_021	2100001	2210000	8800001
AP008216_022	2200001	2310000	8900001
AP008216_023	2300001	2410000	9000001
AP008216_024	2400001	2510000	9100001
AP008216_025	2500001	2610000	9200001
AP008216_026	2600001	2710000	9300001
AP008216_027	2700001	2810000	9400001
AP008216_028	2800001	2910000	9500001
AP008216_029	2900001	3010000	9600001
AP008216_030	3000001	3110000	9700001
AP008216_031	3100001	3210000	9800001
AP008216_032	3200001	3310000	9900001
AP008216_033	3300001	3410000	10000001
AP008216_034	3400001	3510000	10100000
AP008216_035	3500001	3610000	10200001
AP008216_036	3600001	3710000	10300001
AP008216_037	3700001	3810000	10400001
AP008216_038	3800001	3910000	10500001
AP008216_039	3900001	4010000	10600001
AP008216_040	4000001	4110000	10700001
AP008216_041	4100001	4210000	10800001
AP008216_042	4200001	4310000	10900001
AP008216_043	4300001	4410000	11000001
AP008216_044	4400001	4510000	11100001
AP008216_045	4500001	4610000	11200001
AP008216_046	4600001	4710000	11300001
AP008216_047	4700001	4810000	11400001
AP008216_048	4800001	4910000	11500001
AP008216_049	4900001	5010000	11600001
AP008216_050	5000001	5110000	11700001
AP008216_051	5100001	5210000	11800001
AP008216_052	5200001	5310000	11900001
AP008216_053	5300001	5410000	12000001
AP008216_054	5400001	5510000	12100001
AP008216_055	5500001	5610000	12200001
AP008216_056	5600001	5710000	12300001
AP008216_057	5700001	5810000	12400001
AP008216_058	5800001	5910000	12500001
AP008216_059	5900001	6010000	12600001
AP008216_060	6000001	6110000	12700001
AP008216_061	6100001	6210000	12800001
AP008216_062	6200001	6310000	12900001
AP008216_063	6300001	6410000	13000001
AP008216_064	6400001	6510000	13100001
AP008216_065	6500001	6610000	13200001
AP008216_066	6600001	6710000	13300001
AP008216_067	6700001	6810000	13400001

AP008216_068	6800001	6910000
AP008216_069	6900001	7010000
AP008216_070	7000001	7110000
AP008216_071	7100001	7210000
AP008216_072	7200001	7310000
AP008216_073	7300001	7410000
AP008216_074	7400001	7510000
AP008216_075	7500001	7610000
AP008216_076	7600001	7710000
AP008216_077	7700001	7810000
AP008216_078	7800001	7910000
AP008216_079	7900001	8010000
AP008216_080	8000001	8110000
AP008216_081	8100001	8210000
AP008216_082	8200001	8310000
AP008216_083	8300001	8410000
AP008216_084	8400001	8510000
AP008216_085	8500001	8610000
AP008216_086	8600001	8710000
AP008216_087	8700001	8810000
AP008216_088	8800001	8910000
AP008216_089	8900001	9010000
AP008216_090	9000001	9110000
AP008216_091	9100001	9210000
AP008216_092	9200001	9310000
AP008216_093	9300001	9410000
AP008216_094	9400001	9510000
AP008216_095	9500001	9610000
AP008216_096	9600001	9710000
AP008216_097	9700001	9810000
AP008216_098	9800001	9910000
AP008216_099	9900001	10010000
AP008216_100	10000001	10110000
AP008216_101	10100001	10210000
AP008216_102	10200001	10310000
AP008216_103	10300001	10410000
AP008216_104	10400001	10510000
AP008216_105	10500001	10610000
AP008216_106	10600001	10710000
AP008216_107	10700001	10810000
AP008216_108	10800001	10910000
AP008216_109	10900001	11010000
AP008216_110	11000001	11110000
AP008216_111	11100001	11210000
AP008216_112	11200001	11310000
AP008216_113	11300001	11410000
AP008216_114	11400001	11510000
AP008216_115	11500001	11610000
AP008216_116	11600001	11710000
AP008216_117	11700001	11810000
AP008216_118	11800001	11910000
AP008216_119	11900001	12010000
AP008216_120	12000001	12110000
AP008216_121	12100001	12210000
AP008216_122	12200001	12310000
AP008216_123	12300001	12410000
AP008216_124	12400001	12510000
AP008216_125	12500001	12610000
AP008216_126	12600001	12710000
AP008216_127	12700001	12810000
AP008216_128	12800001	12910000
AP008216_129	12900001	13010000
AP008216_130	13000001	13110000
AP008216_131	13100001	13210000
AP008216_132	13200001	13310000
AP008216_133	13300001	13410000
AP008216_134	13400001	13510000
AP008216_135	13500001	13610000
AP008216_136	13600001	13710000
AP008216_137	13700001	13810000
AP008216_138	13800001	13910000
AP008216_139	13900001	14010000
AP008216_140	14000001	14110000

AP008216_141	14100001	14210000	HTG.	Oryza sativa (japonica cultivar-group)
AP008216_142	14200001	14310000	SOURCE	Oryza sativa (japonica cultivar-group)
AP008216_143	14300001	14410000	ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
AP008216_144	14400001	14510000		Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
AP008216_145	14500001	14610000		Ehrhartoideae; Oryzae; Oryza.
AP008216_146	14600001	14710000	REFERENCE	1 (bases 1 to 139999)
AP008216_147	14700001	14810000	AUTHORS	Buell, C.R., Yuan, Q., Moffat, K.S., Hill, J.N., Burr, P.C., Hsiao, J.,
AP008216_148	14800001	14910000		Zismann, V., Pai, G., Bowman, C.L., Fujii, C.Y., VanAken, S.E.,
AP008216_149	14900001	15010000		Bowman, C.L., Craven, B., Utterback, T.R., Khalak, H., Feldblyum, T.V.,
AP008216_150	15000001	15110000		Quackenbush, J., White, O., Salzberg, S.L. and Fraser, C.M.
AP008216_151	15100001	15210000	TITLE	Oryza sativa chromosome 10 BAC OSJNBa0056G17 genomic sequence
AP008216_152	15200001	15310000	JOURNAL	Unpublished
AP008216_153	15300001	15410000	REFERENCE	2 (bases 1 to 139999)
AP008216_154	15400001	15510000	AUTHORS	Buell, R.
AP008216_155	15500001	15610000		Direct Submission
AP008216_156	15600001	15710000	JOURNAL	Submitted (17-DEC-1999) The Institute for Genomic Research, 9712
AP008216_157	15700001	15810000		Medical Center Dr, Rockville, MD 20850, USA
AP008216_158	15800001	15910000	REFERENCE	3 (bases 1 to 139999)
AP008216_159	15900001	16010000	AUTHORS	Buell, R.
AP008216_160	16000001	16110000		Direct Submission
AP008216_161	16100001	16210000	JOURNAL	Submitted (05-JAN-2001) The Institute for Genomic Research, 9712
AP008216_162	16200001	16310000		Medical Center Dr, Rockville, MD 20850, USA, rbuella@tigr.org
AP008216_163	16300001	16410000	REFERENCE	4 (bases 1 to 139999)
AP008216_164	16400001	16510000	AUTHORS	Buell, R.
AP008216_165	16500001	16610000		Direct Submission
AP008216_166	16600001	16710000	JOURNAL	Submitted (13-MAR-2001) The Institute for Genomic Research, 9712
AP008216_167	16700001	16810000		Medical Center Dr, Rockville, MD 20850, USA, rbuella@tigr.org
AP008216_168	16800001	16910000	REFERENCE	5 (bases 1 to 139999)
AP008216_169	16900001	17010000	AUTHORS	Buell, R.
AP008216_170	17000001	17110000		Direct Submission
AP008216_171	17100001	17210000	JOURNAL	Submitted (29-AUG-2001) The Institute for Genomic Research, 9712
AP008216_172	17200001	17310000		Medical Center Dr, Rockville, MD 20850, USA, rbuella@tigr.org
AP008216_173	17300001	17410000	COMMENT	On Jan 5, 2001 this sequence version replaced gi:12025621.
AP008216_174	17400001	17510000		Address all correspondence to:rice@tigr.org
AP008216_175	17500001	17610000		BAC clone OSJNBa0056G17 is from Oryza sativa chromosome 10
AP008216_176	17600001	17710000		The orientation of the sequence is from SP6 to T7 end of the BAC
AP008216_177	17700001	17810000		clone.
AP008216_178	17800001	17910000		Genes were identified by a combination of several methods: Gene
AP008216_179	17900001	18010000		prediction programs including Genscan and Genscan+ (Chris Burge,
AP008216_180	18000001	18110000		http://CCR-081.mit.edu/GENSCAN.html), GeneMarkHM (Mark Borodovsky,
AP008216_181	18100001	18210000		http://genemark.biology.gatech.edu/GeneMark/), Fgenesh
AP008216_182	18200001	18310000		(http://www.softberry.com/), and GeneSplicer (Mihaela Pertea and
AP008216_183	18300001	18410000		Steven Salzberg, contact mpertea@tigr.org), searches of the
AP008216_184	18400001	18510000		complete sequence against a peptide database and the plant EST
AP008216_185	18500001	18610000		database at TIGR (http://www.tigr.org/tdb/cgi.shtml). Annotated
AP008216_186	18600001	18710000		genes are named to indicate the level of evidence for their
AP008216_187	18700001	18810000		annotation. Genes with similarity to other proteins are named after
AP008216_188	18800001	18910000		the database hits. Genes without significant peptide similarity but
AP008216_189	18900001	19010000		with EST similarity are named as unknown proteins. Genes without
AP008216_190	19000001	19110000		protein or EST similarity, that are predicted by more than two gene
AP008216_191	19100001	19210000		prediction programs over most of their length are annotated as
AP008216_192	19200001	19310000		hypothetical proteins. Genes encoding tRNAs are predicted by
AP008216_193	19300001	19410000		hypothetical proteins. Genes encoding tRNAs are predicted by
AP008216_194	19400001	19510000		tRNAscan-SE (Sean Eddy, http://genome.wustl.edu/eddy/tRNAscan-SE/).
AP008216_195	19500001	19610000		Simple repeats are identified by repeatmasker (Arian Smit,
AP008216_196	19600001	19710000		http://ftp.genome.washington.edu/RM/RepeatMasker.html).
Query Match 74.4%; Score 18.6; DB 15; Length 110000;			FEATURES	Location/Qualifiers
Best Local Similarity 84.0%; Pred. No. 5.7e+02;			source	1. .139999
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;				/organism="Oryza sativa (japonica cultivar-group)"
				/mol_type="genomic DNA"
				/cultivar="Nipponbare"
				/sub_species="japonica"
				/db_xref="taxon:39947"
				/chromosome="10"
				/map="C405"
				/clone="OSJNBa0056G17"
			gene	1. .213
				/gene="OSJNBa0056G17.16"
				/notes="similar to cellulose synthase 8 GB:AAF89968
				GI:9622888 (Zea mays)"
			mRNA	<1. .>213
				/gene="OSJNBa0056G17.16"
QY 1 AAAAAAGCACACGAGCCCTCGG 25				
DB 76967 AAAAAAGCACACGACTAGCACTCAG 76943				
RESULT 18				
AC018727/c				
LOCUS				
DEFINITION				
AC018727				
Oryza sativa chromosome 10 BAC OSJNBa0056G17 genomic sequence,				
complete sequence.				
AC018727				
ACCESSION				
AC018727.10				
VERSION				
GI:12039362				

TITLE
JOURNAL
REFERENCE
AUTHORS

Mus musculus chromosome 8, clone RP24-149112
Unpublished
2 (bases 1 to 156342)
Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N.,
Anderson, M., Anderson, S., Arachchi, H.M., Barna, N., Bastien, V.,
Bloom, T., Boguslavskiy, L., Boukhgalter, B., Camarata, J., Chang, J.,
Choepl, Y., Collumore, A., Cook, A., Cooke, P., Corum, B.,
DeArellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L.,
Erickson, J., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D.,
Galgan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N.,
Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I.,
Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T.,
Levine, R., Lindblad-Toh, K., Liu, G., Liu, X., Lui, A., Mabbitt, R.,
MacLean, C., MacDonald, P., Major, J., Manning, J., Matthews, C.,
McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V.,
Murphy, T., Naylor, J., Nguyen, C., Nguyen, T., Nicol, R., Norbu, C.,
O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K.,
Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C.,
Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupback, R.,
Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N.,
Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J.,
Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R.,
Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L.,
Zimmer, A. and Zody, M.

TITLE
JOURNAL
REFERENCE
AUTHORS

Submitted (22-MAR-2005) Broad Institute of MIT and Harvard, 320
Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 156342)
Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N.,
Anderson, M., Anderson, S., Arachchi, H.M., Barna, N., Bastien, V.,
Bloom, T., Boguslavskiy, L., Boukhgalter, B., Camarata, J., Chang, J.,
Choepl, Y., Collumore, A., Cook, A., Cooke, P., Corum, B.,
DeArellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L.,
Erickson, J., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D.,
Galgan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N.,
Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I.,
Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T.,
Levine, R., Lindblad-Toh, K., Liu, G., Liu, X., Lui, A., Mabbitt, R.,
MacLean, C., MacDonald, P., Major, J., Manning, J., Matthews, C.,
McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V.,
Murphy, T., Naylor, J., Nguyen, C., Nguyen, T., Nicol, R., Norbu, C.,
O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K.,
Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C.,
Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupback, R.,
Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N.,
Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J.,
Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R.,
Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L.,
Zimmer, A. and Zody, M.

TITLE
JOURNAL
REFERENCE
AUTHORS

Submitted (02-JUN-2005) Broad Institute of MIT and Harvard, 320
Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 156342)
Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N.,
Anderson, M., Anderson, S., Arachchi, H.M., Barna, N., Bastien, V.,
Bloom, T., Boguslavskiy, L., Boukhgalter, B., Camarata, J., Chang, J.,
Choepl, Y., Collumore, A., Cook, A., Cooke, P., Corum, B.,
DeArellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L.,
Erickson, J., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D.,
Galgan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N.,
Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I.,
Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T.,
Levine, R., Lindblad-Toh, K., Liu, G., Liu, X., Lui, A., Mabbitt, R.,
MacLean, C., MacDonald, P., Major, J., Manning, J., Matthews, C.,
McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V.,
Murphy, T., Naylor, J., Nguyen, C., Nguyen, T., Nicol, R., Norbu, C.,
O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K.,
Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C.,
Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupback, R.,
Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N.,
Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J.,
Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R.,
Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L.,
Zimmer, A. and Zody, M.

TITLE
JOURNAL

Zimmer, A. and Zody, M.
Direct Submission
Submitted (28-JUL-2005) Broad Institute of MIT and Harvard, 320
Charles Street, Cambridge, MA 02141, USA

COMMENT

On Jul 28, 2005 this sequence version replaced gi:66865067.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center
Center: Broad Institute of MIT and Harvard

Center code: W1BR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submission@broad.mit.edu
----- Project Information
Center project name: L31955
Center clone name: 149_I_12

FEATURES
source

----- Location/Qualifiers
1..156342
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="8"
/map="8"
/clone="RP24-149112"
322..2360 /rpt_family="AT_rich"
4476..4499 /rpt_family="AT_rich"
5021..5061 /rpt_family="AT_rich"
complement(5158..6310)
/rpt_family="RMR12"
6341..6404
/rpt_family="MTB"
6405..6687 /rpt_family="Lx8"
7121..8395 /rpt_family="Lx8"
/rpt_family="Lx8"
8390..14584
/rpt_family="L1_MM"
14592..15204
/rpt_family="L1"
15191..16139
/rpt_family="L1"
16142..16184
/rpt_family="AT_rich"
16740..16927
/rpt_family="B2_Mm2"
18503..18546
/rpt_family="(CA)n"
19421..19575
/rpt_family="B3A"
19605..21143
/rpt_family="L1_MM"
21265..21515
/rpt_family="L1ME"
22058..22262
/rpt_family="B3"
22724..22800
/rpt_family="AT_rich"
complement(23142..23259)
/rpt_family="ID_B1"
complement(23346..23635)
/rpt_family="Lx2"
24022..24064
/rpt_family="AT_rich"
24238..24315
/rpt_family="MYSERV"
24740..24783
/rpt_family="(TCC)n"
24905..24957

```

/rpt family=" (TC)n"
24958. .25002
/rpt family=" (CA)n"
25058. .25502
/rpt family="RNER20"
26109. .26257
/rpt family="RSINE1"
26280. .26330
/rpt family=" (CA)n"
complement(27490. .27826)
/rpt family="IAPLTR1_MM"
complement(27827. .32869)
/rpt family="IAPLTR1_MM-int"
complement(32870. .33012)
/rpt family="MUT1E1A"
complement(33013. .34208)
/rpt family="IAPLTR1_MM-int"
complement(34209. .34545)
/rpt family="IAPLTR1_MM"
34745. .34928
/rpt family="L1"
35655. .35683
/rpt family=" (GAAA)n"
35686. .35741
/rpt family=" (GGGA)n"
36660. .36822
/rpt family=" (TCC)n"
37783. .38108
/rpt family="ORR1A2"
complement(38140. .38578)
/rpt family="L1_MM"
38909. .39103
/rpt family="B3"
39291. .39319
/rpt family="AT_rich"
39505. .39543
/rpt family=" (TCTCTG)n"
complement(39590. .39665)
/rpt family="B4A"
39820. .40214
/rpt family="MTA"
40410. .40486

Query Match 74.4%; Score 18.6; DB 9; Length 156342;
Best Local Similarity 84.0%; Pred. No. 5.6e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAACGACACAGCGAGCCCTCGG 25
Db 32414 AAAAAACGAATCATCGGCGCCTCGG 32390

RESULT 20
AC160549
LOCUS AC160549 157085 bp DNA linear HTG 10-JUN-2005
DEFINITION Mus musculus chromosome 3 clone RP23-276K18 map 3, *** SEQUENCING
IN PROGRESS ***, 2 ordered pieces.
ACCESSION AC160549
VERSION AC160549.4 GI:67460011
KEYWORDS HTG; HTGS_PHASE2; HTGS_FULLTOP; HTGS_ACTIVEFIN.
SOURCE Mus musculus
ORGANISM Mus musculus (house mouse)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 157085)
REFERENCE
AUTHORS Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
TITLE Mus musculus chromosome 3, Clone RP23-276K18
JOURNAL Unpublished
REFERENCE
AUTHORS Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
Anderson,M., Anderson,S., Atachchi,H.M., Barna,N., Bastien,V.,
Bloom,T., Boguslavskiy,L., Boukhgalter,B., Camarata,J., Chang,J.,
Choepl,Y., Collymore,A., Cook,A., Cooke,P., Corum,B.,
DeArellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L.,
Erickson,J., Fato,S., Ferreira,P., FitzGerald,M., Gage,D.,
Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N.,
Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I.,
Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T.,
Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbitt,R.,
MacLean,C., Macdonald,P., Major,J., Manning,J., Matthews,C.,
McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
Murphy,T., Naylor,J., Nguyen,C., Nguyen,T., Nicol,R., Norbu,C.,
O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,J., Raymond,C.,
Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R.,
Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L.,
Zimmer,A. and Zody,M.
Center: Broad Institute of MIT and Harvard
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@broad.mit.edu
----- Project Information
Center project name: L32177
Center clone name: 276_K_18
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 115320: contig of 115320 bp in length
* 115321 115420: gap of unknown length
* 115421 157085: contig of 41665 bp in length.
* Location/Qualifiers
1. .157085
source
```

```

/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="4"
/clone="RP23-86013"
/clone_lib="RPCI-23"
/map="3"
/clone="RP23-276K18"
/clone_lib="RPCI-23 Female Mouse BAC"
115321..115420
/estimated_length=unknown

gap
Query Match 74.4%; Score 18.6; DB 14; Length 157085;
Best Local Similarity 84.0%; Pred. No. 5.6e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

ORIGIN
QY 1 AAAAAACGACACGAGCCCTCGG 25
   ||||| ||||| ||||| ||||| |||||
Db 72877 AAAAAACGCGCCATCGAGCCCTCGG 72901

RESULT 21
AL805943 192561 bp DNA linear ROD 17-DEC-2004
LOCUS Mouse DNA sequence from clone RP23-86013 on chromosome 4 Contains a
DEFINITION opsin 1 (cone pigments) (Opnl) pseudogene, complete sequence.
ACCESSION AL805943
VERSION AL805943.6 GI:46406414
KEYWORDS HTG; Opnl; opsin.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 192561)
AUTHORS North,P., Leaves,N., Greystrom,J., Coppola,M., Manjunath,S.,
Russell,E., Smith,M., Strachan,G., Tofts,C., Boal,E., Cobley,V.,
Hunter,G., Kimberley,C., Thomas,D., Cave-Berry,L., Weston,P. and
Botcherby,M.R.M.
Direct Submission
Submitted (16-DEC-2004) Mouse Sequencing Group, HGMP-RC, Hinxton,
Cambridge, CB10 1SB, UK. E-mail enquiries:- mrbotche@hgmp.mrc.ac.uk
or phorthehgmp.mrc.ac.uk
HGMPC-RC part of the UK Mouse Sequencing Consortium
On Apr 19, 2004 this sequence version replaced gi:28172099.
----- Genome Center
Center: UK Medical Research Council
Center code: UK-MRC
Web site: http://mrcseq.har.mrc.ac.uk
Contact: mouseq@har.mrc.ac.uk
-----
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep/RP23-86013 is
from the RPCI-23 Mouse BAC Library
constructed by the group of Pieter de Jong.
For further details see http://www.chori.org/bacpac/home.htm
VECTOR: pBACe3.6
Sequence from the Mouse Genome Sequencing Consortium whole genome
shotgun may have been used to confirm this sequence. Sequence data
from the whole genome shotgun alone has only been used where it has
a phred quality of at least 30.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one subclone; and the assembly was confirmed by restriction digest,
except on the rare occasion of the clone being a YAC.
FEATURES
source
1..192561
/organism="Mus musculus"

/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="4"
/clone="RP23-86013"
/clone_lib="RPCI-23"
175628..176030
/locus_tag="RP23-86013.1-001"
/pseudo
175628..176030
/locus_tag="RP23-86013.1-001"
/notes="match: Proteins: SW:Q35599 Sw:Q9W6A7 Tr:Q9JLS7 Tr:Q9QY62
Tr:Q864C4 Tr:Q864C9 Tr:Q98UH7 Tr:Q9JLS7 Tr:Q9QY62
Tr:Q9Y152"
/codon_start=1
/product="opsin 1 (cone pigments) (Opnl) pseudogene"

ORIGIN
Query Match 74.4%; Score 18.6; DB 9; Length 192561;
Best Local Similarity 84.0%; Pred. No. 5.5e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAACGACACGAGCCCTCGG 25
   ||||| ||||| ||||| ||||| |||||
Db 191609 AAAAAATGACCCAGCTAGCCCTCTG 191633

RESULT 22
AC122175 196097 bp DNA linear PRI 18-SEP-2003
LOCUS Pan troglodytes clone rp43-99g12, complete sequence.
DEFINITION AC122175
ACCESSION AC122175
VERSION AC122175.9 GI:34849974
KEYWORDS HTG.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
REFERENCE 1 (bases 1 to 196097)
AUTHORS Shi,R., Fu,Y., Zhou,L., Wu,J., Shaull,S., Eichler,E. and Roe,B.A.
TITLE Pan troglodytes BAC Clone rp43-99g12
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 196097)
AUTHORS Shi,R., Fu,Y., Zhou,L., Wu,J., Shaull,S., Eichler,E. and Roe,B.A.
TITLE Direct Submission
JOURNAL Submitted (23-MAY-2002) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
REFERENCE 3 (bases 1 to 196097)
AUTHORS Shi,R., Fu,Y., Zhou,L., Wu,J., Shaull,S., Eichler,E. and Roe,B.A.
TITLE Direct Submission
JOURNAL Submitted (02-JUL-2002) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
REFERENCE 4 (bases 1 to 196097)
AUTHORS Shi,R., Fu,Y., Zhou,L., Wu,J., Shaull,S., Eichler,E. and Roe,B.A.
TITLE Direct Submission
JOURNAL Submitted (29-JUL-2002) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
REFERENCE 5 (bases 1 to 196097)
AUTHORS Shi,R., Fu,Y., Zhou,L., Wu,J., Shaull,S., Eichler,E. and Roe,B.A.
TITLE Direct Submission
JOURNAL Submitted (22-OCT-2002) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
REFERENCE 6 (bases 1 to 196097)
AUTHORS Shi,R., Fu,Y., Zhou,L., Wu,J., Shaull,S., Eichler,E. and Roe,B.A.
TITLE Direct Submission
JOURNAL Submitted (23-OCT-2002) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA

```

```

REFERENCE
AUTHORS
TITLE
JOURNAL

8 (bases 1 to 196097)
Shi R., Fu Y., Zhou L., Wu J., Shaul L., Shaul S., Eichler E. and Roe B.A.
Direct Submission
Submitted (23-JAN-2003) Department of Chemistry And Biochemistry,
The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA

REFERENCE
AUTHORS
TITLE
JOURNAL

8 (bases 1 to 196097)
Shi R., Fu Y., Zhou L., Wu J., Shaul L., Shaul S., Eichler E. and Roe B.A.
Direct Submission
Submitted (18-SEP-2003) Department of Chemistry And Biochemistry,
The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA

COMMENT
On Sep 18, 2003 this sequence version replaced gi:27877221.

----- Genome Center
Center: Department Of Chemistry And Biochemistry
The University Of Oklahoma
Center code:UOKNOR

-----
Location/Qualifiers
1. .196097
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clone="rp43-99g12"
/clone_lib="RPCI - 43 Male Chimpanzee BAC Library"

ORIGIN
Query Match 74.4%; Score 18.6; DB 8; Length 196097;
Best Local Similarity 84.0%; Pred. No. 5.5e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAAGCAGACAGCGAGCCCTCGG 25
||||| ||||| ||||| |||||
Db 152949 AAAAAAGCAGACAGCGAGCCCTCGG 152973

RESULT 23
AC162627
LOCUS Bos taurus clone CH240-124D18, *** SEQUENCING IN PROGRESS ***, 18
DEFINITION AC162627
unordered pieces.
AC162627 2 GI:68301357
HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE Bos taurus
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
Pecora; Bovidae; Bovinae; Bos.

1 (bases 1 to 214864)
Muzny D., Marie M., Metzker M., Lee A., Abramson S., Adams C., Alder J.,
Allen C., Allen H., Alsbrooks S., Amin A., Anguiano D.,
Anyalebechi V., Aoyagi A., Ayodeji M., Baca E., Baden H.,
Baldwin D., Bandaranaike D., Barber M., Barnstead M., Benahmed F.,
Biswal O., Blair J., Blankenburg K., Blyth P., Brown M.,
Bryant N., Buhay C., Burch P., Burrell K., Calderon E.,
Cardenas V., Carter K., Cavazos I., Ceasar H., Center A.,
Chacko J., Chavez D., Chen G., Chen R., Chen Y., Chen Z., Chu J.,
Cleveland C., Cockrell R., Cox C., Coyle M., Cree A., D'Souza L.,
Davila M., Davis C., Davy-Carroll L., De Anda C., Dederich D.,
Delgado O., Denson S., Deramo C., Ding Y., Dinh H., Divya K.,
Draper H., Dugan-Rocha S., Dunn A., Durbin K., Duval B., Eaves K.,
Egan A., Escotto M., Eugene C., Evans C.A., Falls T., Fan G.,
Fernandez S., Finley M., Flagg N., Forbes L., Foster M., Foster P.,
Fraser C.M., Gabisi A., Ganta R., Garcia A., Garner T., Garza M.,
Gebrgeorgis E., Geer K., Gill R., Grady M., Guerra W., Guevara W.,
Gunaratne P., Haaland W., Hamil C., Hamilton C., Hamilton K.,
Harvey Y., Havlak P., Haves A., Henderson N., Hernandez J.,
Hernandez R., Hines S., Hladun S.L., Hodgson A., Hogues M.,
Hollins B., Howells S., Hulyk S., Hume J., Idlebird D., Jackson A.,
Jackson L., Jacob L., Jiang H., Johnson B., Johnson R., Jolivet A.,
Karpachy S., Kelly S., Kelly S., Khan Z., King L., Kovar C.,
Kowis C., Kraft C.L., Lebow H., Levan J., Lewis L., Li Z., Liu J.,

```

```

Liu J., Liu W., Liu Y., London P., Longacre S., Lopez J.,
Lorensuheva L., Loulseged H., Lozado R.J., Lu X., Ma J.,
Maheshwari M., Mahindartne M., Mahmoud M., Malloy K., Mangum A.,
Mangum B., Mapua P., Martin K., Martin R., Martinez E.,
Mahoney S., McLeod M.P., McNeill T.Z., Meenen E.,
Milosavljevic A., Miner G., Minja E., Montemayor J., Moore S.,
Morgan M., Morris K., Morris S., Munidasa M., Murphy M., Nair L.,
Nankervis C., Neal D., Newton N., Nguyen N., Norris S.,
Nwaokemele O., Okwuonu G., Olarnpunsagoon A., Pal S., Parks K.,
Pasternak S., Paul H., Perez A., Perez L., Pfannkuch C.,
Plummer F., Poindexter A., Popovic D., Primus E., Pu L.-L.,
Puzo M., Quiroz J., Rachlin E., Reeves K., Regier M.A., Reigh R.,
Reilly B., Reilly M., Ren Y., Reuter M., Richards S., Riggs F.,
Rives C., Rodkey T., Rojas A., Rose M., Rose R., Ruiz S.J.,
Sanders W., Savery G., Scherer S., Scott G., Shatman S., Shen H.,
Shetty J., Shivabtsbeyn A., Sisson I., Sitter C.D., Smajd D.,
Sneid A., Sodergren E., Song X.-Z., Sorelle R., Sosai J.,
Steinle M., Strong R., Sutton A., Svatek A., Tabor P., Taylor C.,
Taylor T., Thomas N., Thomas S., Tingey A., Trejos Z., Umani K.,
Valas R., Vera V., Villasana D., Waldron L., Walker B., Wang J.,
Wang Q., Wang S., Warren J., Warren R., Wei X., White F.,
Williams G., Willson R., Wleczek R., Wooden H., Worley K.,
Wright D., Wright R., Wu J., Yakub S., Yen J., Yoon L., Yoon V.,
Yu F., Zhang J., Zhou J., Zhou X., Zhao S., Dunn D., von
Niederhausern A., Weiss R., Smith D.R., Holt R.A., Smith H.O.,
Weinstock G. and Gibbs R.A.

Direct Submission
Unpublished
2 (bases 1 to 214864)
Worley K.C.
Direct Submission
Submitted (31-MAY-2005) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 214864)
Baylor Genome Sequencing Consortium.
Direct Submission
Submitted (01-JUL-2005) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jun 29, 2005 this sequence version replaced gi:66796257.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: FHNA
Center clone name: CH240-124D18
----- Summary Statistics
Assembly program: Atlas 3.0;
Consensus quality: 189868 bases at least Q40
Consensus quality: 191139 bases at least Q30
Consensus quality: 192213 bases at least Q20
Estimated insert size: 194905; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 18 contigs. The true order of the pieces

```


McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V.,
 Murphy, T., Naylor, J., Nguyen, C., Nguyen, T., Nicol, R., Norbu, C.,
 O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K.,
 Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C.,
 Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schuback, R.,
 Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N.,
 Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J.,
 Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R.,
 Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L.,
 Zimmer, A. and Zody, M.

TITLE
JOURNAL
COMMENT
 Submitted (23-JUL-2005) Broad Institute of MIT and Harvard, 320
 Charles Street, Cambridge, MA 02141, USA
 On Jul 23, 2005 this sequence version replaced gi:68989329.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html
 ----- Genome Center
 Center: Broad Institute of MIT and Harvard
 Center code: WIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@broad.mit.edu
 ----- Project Information
 Center project name: L30779
 Center clone name: 420_P_19

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 3 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

* 1 37065: contig of 37065 bp in length
 * 37066 37165: gap of unknown length
 * 37166 174695: contig of 137530 bp in length
 * 174696 174795: gap of unknown length
 * 174796 218219: contig of 43424 bp in length.

FEATURES

source

1. 218219
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /chromosomes="1"
 /map="1"
 /clone="RP23-420P19"
 /clone_lib="RP23-420P19"
 /clone="RP23-420P19"
 37066..37165
 /estimated_length=unknown
 174696..174795
 /estimated_length=unknown

gap

gap

ORIGIN

Query Match 74.4%; Score 18.6; DB 14; Length 218219;
 Best Local Similarity 84.0%; Pred. No. 5.5e+02;
 Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAACGACACAGCGAGCCCTCGG 25

Db 169680 AAAAAACGAGTCATCGGCGCCCTCGG 169704

RESULT 25

AC115721 218746 bp DNA linear ROD 29-OCT-2003
LOCUS
 AC115721 Mus musculus chromosome 3, clone RP23-9p8, complete sequence.
DEFINITION
 AC115721
ACCESSION
 AC115721.11 GI:38016090
VERSION
 HTG
KEYWORDS
 Mus musculus (house mouse)
SOURCE
 Mus musculus
ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Murioidea; Muridae; Murinae; Mus.
 1 (bases 1 to 218746)
 Birren, B., Nussbaum, C. and Lander, E.
 Mus musculus chromosome 3, clone RP23-9p8
 Unpublished
 2 (bases 1 to 218746)
 Birren, B., Linton, L., Nussbaum, C., Lander, E., Ali, A., Allen, N.,
 Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L.,
 Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J.,
 Chazaro, B., Choepel, Y., Collangelo, M., Collins, S., Collymore, A.,
 Cook, A., Cooke, P., DeArelano, K., Dewar, K., Diaz, J.S., Dodge, S.,
 Fero, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S.,
 Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
 Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
 Kamat, A., Karatas, A., Kells, C., LaRoque, K., Lamarcares, R.,
 Landers, T., Lehoczy, J., Levine, R., Lindblad-Toh, K., Liu, G.,
 MacLean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C.,
 McCarthy, M., McEwan, P., McKernan, K., Meldrim, J., Meneus, L.,
 Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R.,
 Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D.,
 Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,
 Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
 Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schuback, R.,
 Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
 Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
 Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,
 Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,
 Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission

Submitted (22-MAR-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA

3 (bases 1 to 218746)

REFERENCE
AUTHORS

Birren, B., Nussbaum, C., Lander, E., Abouelleil, A., Allen, N.,
 Anderson, M., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T.,
 Boguslavskiy, L., Boukhgalter, B., Camarata, J., Chang, J., Choepel, Y.,
 Collymore, A., Cook, A., Cooke, P., Corum, B., DeArelano, K.,
 Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Fero, S.,
 Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S.,
 Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B.,
 Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
 Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R.,
 Lindblad-Toh, K., Liu, G., Lui, X., Mabbitt, R., MacLean, C.,
 Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M.,
 Meldrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J.,
 Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P.,
 O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N.,
 Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P.,
 Roman, J., Schauer, S., Schuback, R., Seaman, S., Severy, P., Smith, C.,
 Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M.,
 Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M.,
 Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X.,
 Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission

Submitted (05-OCT-2003) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA

4 (bases 1 to 218746)

REFERENCE
AUTHORS

Birren, B., Nussbaum, C., Lander, E., Abouelleil, A., Allen, N.,
 Anderson, M., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T.,
 Boguslavskiy, L., Boukhgalter, B., Camarata, J., Chang, J., Choepel, Y.,
 Collymore, A., Cook, A., Cooke, P., Corum, B., DeArelano, K.,
 Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Fero, S.,
 Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S.,
 Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B.,
 Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
 Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R.,
 Lindblad-Toh, K., Liu, G., Lui, X., Mabbitt, R., MacLean, C.,
 Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M.,
 Meldrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J.,
 Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P.,
 O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N.,
 Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P.,
 Roman, J., Schauer, S., Schuback, R., Seaman, S., Severy, P., Smith, C.,
 Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M.,
 Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M.,
 Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X.,
 Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M.,
 Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M.,
 Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X.,
 Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
 Direct Submission
 Submitted (29-Oct-2003) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Oct 29, 2003 this sequence version replaced gi:37515300.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html
 ----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WITBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L22279
 Center clone name: 9_p_8

FEATURES

source Location/Qualifiers
 1. .218746
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /chromosome="3"
 /map="3"
 /clone="RP23-9P8"
 /clone_lib="RPC1-23 Female Mouse BAC"
 1. .459
 /rpt_family="MWERVL_LTR"
 1. .143
 /note="wgs end extension
 clone_end:SP6"
 /note="single clone coverage"
 complement(39. .107)
 /note="single clone coverage"
 complement(68. .80)
 /note="<30 qual single clone coverage"
 144. .149
 /note="clone boundary
 clone_end:SP6
 site:ECORI"
 complement(732. .6710)
 /rpt_family="L1P"
 complement(4845)
 /note="some subclones have A, not G"
 complement(5614. .5642)
 /note="single clone coverage"
 complement(6711. .6949)
 /rpt_family="L1P"
 6957. .7116
 /rpt_family="Tigger7"
 8101. .8205
 /rpt_family="MTD"
 8572. .8700
 /rpt_family="Lx8"
 complement(10095. .10273)
 /rpt_family="B2_Mm2"
 10306. .10348
 /rpt_family="T-rich"
 complement(10354. .10476)
 /rpt_family="B1_MM"
 10627. .10657
 /rpt_family="T)n"
 10943. .10998
 /rpt_family="CT-rich"
 complement(10999. .11116)
 /rpt_family="B1_MM"
 complement(11145. .11216)
 /rpt_family="ID_B1"
 complement(11498. .11622)
 /rpt_family="B1F"
 14610. .14793

repeat_region
 complement(15055. .15243)
 /rpt_family="B3A"
 15308. .15340
 /rpt_family="AT_rich"
 complement(15720. .15856)
 /rpt_family="B1_MM"
 complement(15728. .15922)
 /rpt_family="ID_B1"
 16683. .16948
 /rpt_family="B4A"
 17330. .17731
 /rpt_family="RLTR9B"
 17785. .17935
 /rpt_family="RSINE1"
 17982. .18192
 /rpt_family="B3"
 18456. .18753
 /rpt_family="LIM3"
 19692. .19898
 /rpt_family="B3"
 20258. .20336
 /rpt_family="B1F"
 20337. .20358
 /rpt_family="ID_B1"
 21130. .21160
 /rpt_family="AT_rich"
 21668. .21715
 /rpt_family="TCTA)n"
 21716. .21738
 /rpt_family="CTTA)n"
 complement(21789. .21901)
 /rpt_family="B4"
 23083. .23111
 /rpt_family="(CA)n"
 23138. .23377
 /rpt_family="LIM4"
 24421. .25333
 /rpt_family="LIM4"
 25328. .31558
 /rpt_family="L1_MM"
 31631. .31973
 /rpt_family="MTC"
 32046. .32635
 /rpt_family="LIM4"
 32636. .32832
 repeat_region

Query Match 74.4%; Score 18.6; DB 9; Length 218746;
 Best Local Similarity 84.0%; Pred. No. 5.5e+02;
 Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAACGACAGCGAGCCCTCGG 25
 |||||
 Db 176526 AAAAAACGAGCCATCGAGCCCTCGG 176550

RESULT 26

AC138614
 LOCUS 230451 bp DNA linear HTG 25-FEB-2003
 DEFINITION Mus musculus chromosome UNK clone RP23-421A7, WORKING DRAFT
 SEQUENCE, 5 unordered pieces.
 AC138614
 ACCESSION AC138614.2 GI:27777615
 VERSION HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
 KEYWORDS Mus musculus (house mouse)
 SOURCE
 ORGANISM
 Mus musculus
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muroidae; Muridae; Murinae; Mus.
 REFERENCE
 1 (bases 1 to 230451)
 AUTHORS McPherson, J.D. and Waterston, R.H.
 TITLE The sequence of Mus musculus clone
 JOURNAL Unpublished

```
REFERENCE
AUTHORS      McPherson,J.D. and Waterston,R.H.
TITLE        Direct Submission
JOURNAL      Submitted (12-JAN-2003) Genome Sequencing Center, 4444 Forest Park
              Parkway, St. Louis, MO 63108, USA
REFERENCE
AUTHORS      McPherson,J.D. and Waterston,R.H.
TITLE        Direct Submission
JOURNAL      Submitted (25-FEB-2003) Genome Sequencing Center, 4444 Forest Park
              Parkway, St. Louis, MO 63108, USA
COMMENT      On Jan 17, 2003 this sequence version replaced gi:27657600.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site:http://genome.wustl.edu/gsc/index.shtml
Contact: submissions@wustl.wustl.edu
----- Project Information -----
Center project name: M BA0421A07
----- Summary Statistics -----
Sequencing vector: M13; 0%
Sequencing vector: plasmid; 100%
Chemistry: Dye-primer ET; 0% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 228476 bases at least Q40
Consensus quality: 228826 bases at least Q30
Consensus quality: 229089 bases at least Q20
Insert size: 202000; agarose-fp
Insert size: 235563; sum-of-contigs
Quality coverage: 11.98 in Q20 bases; agarose-fp
Quality coverage: 11.18 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 5 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1343: contig of 1343 bp in length
* 1344: 1443: gap of unknown length
* 1444: 2955: contig of 1512 bp in length
* 2956: 3055: gap of unknown length
* 3056: 26931: contig of 23876 bp in length
* 26932: 27031: gap of unknown length
* 27032: 87168: contig of 60137 bp in length
* 87169: 87268: gap of unknown length
* 87269: 230451: contig of 143183 bp in length.

FEATURES
source
1..230451
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="UNK"
/clone="RP23-421A7"
1..1343
/notes="assembly_name:Contig6"
1344..1443
/estimated_length=unknown
1444..2955
/notes="assembly_name:Contig21"
2956..3055
/estimated_length=unknown
3056..26931
/notes="assembly_name:Contig22"
26932..27031
/estimated_length=unknown
27032..87168
/notes="assembly_name:Contig23"
87169..87268

-----
Query Match      74.4%; Score 18.6; DB 14; Length 230451;
Best Local Similarity 84.0%; Pred. No. 5.5e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

ORIGIN
1 AAAAAACGACACAGCGAGCCCTCGG 25
202023 AAAAAACGATCATCGGCGCCCTCGG 202047

-----
RESULT 27
AC119312/c
LOCUS      Rattus norvegicus clone CH230-342E8, WORKING DRAFT SEQUENCE.
DEFINITION AC119312
ACCESSION  AC119312
VERSION    AC119312.5 GI:24635493
KEYWORDS   HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE     Rattus norvegicus (Norway rat)
ORGANISM   Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muridea; Muridae; Rattus.
REFERENCE  1 (bases 1 to 236746)
AUTHORS   Muzny,D.Marie., Metzker,M.Lee., Abranzon,S., Adams,C., Alder,J.,
            Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
            Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
            Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
            Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
            Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
            Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
            Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
            Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
            Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
            Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
            Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
            Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
            Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
            Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
            Gebregeorgis,B., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
            Gunaratne,P., Haaland,W., Hamill,C., Hamilton,C., Hamilton,K.,
            Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
            Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hognes,M.,
            Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
            Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
            Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
            Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
            Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
            Lorensenwa,L., Loulsegged,H., Lozado,R.J., Lu,X., Ma,J.,
            Maheshwari,M., Mahindartne,M., Mahmoud,M., Mallory,K., Mangum,A.,
            Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
            Mawhinney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
            Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
            Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
            Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Parks,K.,
            Nwaokemele,O., Okwuonu,G., Olarnpusagoon,A., Pal,S., Parks,K.,
            Pasternak,S., Paul,H., Perez,A., Perez,L., Pfanckoch,C.,
            Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.,
            Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
            Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
            Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
            Sanders,W., Savarysbeyn,A., Sisson,I., Sitter,C.D., Smajls,D.,
            Shetty,J., Shvartsbeyn,A., Sison,I., Scott,G., Shatman,S., Shen,H.,
            Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
            Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
            Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Umani,K.,
            Valas,R., Vera,V., Villanana,D., Waldron,L., Walker,B., Wang,J.,
            Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
            Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
            Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
```

Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
Niederhausern, A., Welles, R., Smith, D.R., Holt, R.A., Smith, H.O.,
Weinstock, G. and Gibbs, R.A.
Direct Submission
Unpublished
2 (bases 1 to 236746)
Worley, K.C.
Direct Submission
Submitted (26-APR-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 236746)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (19-NOV-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Nov 6, 2002 this sequence version replaced gi:23811730.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.
----- Genome Center
Center: Baylor College of Medicine
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GUOA
Center clone name: CH230-342E8
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 172218 bases at least Q40
Consensus quality: 174486 bases at least Q30
Consensus quality: 175960 bases at least Q20
Estimated insert size: 174258; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 236746: contig of 236746 bp in length.

FEATURES
source
1..236746
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-342E8"
1..1128
/note="wgs_end_extension
clone_end:Sp6"
54952..56001
/note="wgs_end_extension
clone_end:Sp6"
56052..58438
/note="wgs_end_extension
clone_end:Sp6"

misc_feature
58489..59494
/note="wgs_end_extension
clone_end:Sp6"
62288..63062
/note="clone boundary
clone_end:Sp6
site:
end_sequence:B2209427"

ORIGIN
Query Match 74.4%; Score 18.6; DB 14; Length 236746;
Best Local Similarity 84.0%; Pred. No. 5.5e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 AAAAAACGACACAGCGAGCCCTCGG 25
|||||
Db 229428 AAAAAACGACACAGCGAGCCCTCAG 229404
|||||

RESULT 28
CR376791
LOCUS
DEFINITION
Mus musculus chromosome 4 clone RP23-30108, *** SEQUENCING IN
PROGRESS ***, 36 unordered pieces.
ACCESSION
CR376791
VERSION
CR376791.1 GI:45598591
KEYWORDS
HTG; HTGS; PHASE1.
SOURCE
Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Murinae; Mus.
1 (bases 1 to 258567)
Burton, J.
Direct Submission
Submitted (01-APR-2004) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
Sequence from the Mouse Genome Sequencing Consortium whole genome
shotgun may have been used to confirm this sequence. Sequence data
from the whole genome shotgun alone has only been used where it has
a phred quality of at least 30.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: <http://www.sanger.ac.uk>
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: BM30108
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; 100% of reads
Consensus quality: 238991 bases at least Q40
Consensus quality: 245718 bases at least Q30
Consensus quality: 250137 bases at least Q20
Insert size: 255067; sum-of-contigs
Insert size: 97305; 5.0% error; agarose-fp
Quality coverage: 3.29x in Q20 bases; sum-of-contigs Quality
coverage: 9.13x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently
* consists of 36 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 15346: contig of 15346 bp in length
* 15347 15446: gap of 100 bp
* 15447 43658: contig of 28212 bp in length
* 43659 43758: gap of 100 bp
* 43759 64450: contig of 20692 bp in length

* 64451	64550: gap of 100 bp		/clone_lib="RPCI-23"
* 64551	67501: contig of 3058 bp in length	misc_feature	1. .15346
* 67609	67708: gap of 100 bp		/note="assembly fragment:00842
* 67709	73421: contig of 6234 bp in length		fragment_chain:1"
* 73943	74042: gap of 100 bp	misc_feature	15447. .43658
* 74043	91996: contig of 17954 bp in length		/note="assembly fragment:01132
* 91997	92096: gap of 100 bp		fragment_chain:1"
* 92097	94347: contig of 2251 bp in length	misc_feature	43759. .64450
* 94348	94447: gap of 100 bp		/note="assembly fragment:00962
* 94448	104001: contig of 9554 bp in length		fragment_chain:1"
* 104002	104101: gap of 100 bp	misc_feature	64551. .67608
* 104102	109235: contig of 5134 bp in length		/note="assembly fragment:00299
* 109236	109335: gap of 100 bp		fragment_chain:1"
* 109336	114190: contig of 4855 bp in length	misc_feature	67709. .73942
* 114191	114290: gap of 100 bp		/note="assembly fragment:00478
* 114291	124963: contig of 10673 bp in length		fragment_chain:1"
* 124964	125063: gap of 100 bp	misc_feature	74043. .91996
* 125064	129213: contig of 4150 bp in length		/note="assembly fragment:00728
* 129214	129313: gap of 100 bp		fragment_chain:1"
* 129314	132069: contig of 2756 bp in length	misc_feature	92097. .94347
* 132070	132169: gap of 100 bp		/note="assembly fragment:00061
* 132170	136359: contig of 4190 bp in length		fragment_chain:2"
* 136360	136459: gap of 100 bp	misc_feature	94448. .104001
* 136460	144197: contig of 7738 bp in length		/note="assembly fragment:00377
* 144198	144297: gap of 100 bp		fragment_chain:2"
* 144298	146301: contig of 2004 bp in length	misc_feature	104102. .109235
* 146302	146401: gap of 100 bp		/note="assembly fragment:00151
* 146403	158833: contig of 12432 bp in length		fragment_chain:2"
* 158834	159333: gap of 100 bp	misc_feature	109336. .114190
* 159334	166350: contig of 7417 bp in length		/note="assembly fragment:00138
* 166351	166450: gap of 100 bp		fragment_chain:2"
* 166451	168700: contig of 2250 bp in length	misc_feature	114291. .124963
* 168701	169800: gap of 100 bp		/note="assembly fragment:00524
* 169801	171153: contig of 2353 bp in length		fragment_chain:2"
* 171154	171253: gap of 100 bp	misc_feature	125064. .129213
* 171254	174131: contig of 2878 bp in length		/note="assembly fragment:00257
* 174132	174231: gap of 100 bp		fragment_chain:3"
* 174232	176746: contig of 2515 bp in length	misc_feature	129314. .132069
* 176747	176846: gap of 100 bp		/note="assembly fragment:00077
* 176847	195914: contig of 19068 bp in length		fragment_chain:3"
* 195915	196014: gap of 100 bp	misc_feature	132170. .136359
* 196015	198745: contig of 2731 bp in length		/note="assembly fragment:00192
* 198746	198845: gap of 100 bp		fragment_chain:3"
* 198846	200865: contig of 2020 bp in length	misc_feature	136460. .144197
* 200866	200965: gap of 100 bp		/note="assembly fragment:00276
* 200966	203893: contig of 2928 bp in length		fragment_chain:3"
* 203894	208244: contig of 4251 bp in length	misc_feature	144298. .146301
* 208245	208344: gap of 100 bp		/note="assembly fragment:00069
* 208346	213055: contig of 4711 bp in length		fragment_chain:4"
* 213056	213155: gap of 100 bp	misc_feature	146402. .158833
* 213156	218757: contig of 5602 bp in length		/note="assembly fragment:00573
* 218758	218857: gap of 100 bp		fragment_chain:4"
* 218858	222834: contig of 3977 bp in length	misc_feature	158934. .166350
* 222835	222934: gap of 100 bp		/note="assembly fragment:00322
* 222935	228500: contig of 5566 bp in length		fragment_chain:4"
* 228501	228600: gap of 100 bp	misc_feature	166451. .168700
* 228601	233666: contig of 5066 bp in length		/note="assembly fragment:00043
* 233667	233766: gap of 100 bp		fragment_chain:5"
* 233767	242489: contig of 8723 bp in length	misc_feature	168801. .171153
* 242490	242589: gap of 100 bp		/note="assembly fragment:00115
* 242590	244863: contig of 2274 bp in length		fragment_chain:5"
* 244864	244963: gap of 100 bp	misc_feature	171254. .174131
* 244964	248270: contig of 3307 bp in length		/note="assembly fragment:00086
* 248271	248370: gap of 100 bp		fragment_chain:6"
* 248371	258567: contig of 10197 bp in length.	misc_feature	174232. .176746
	Location/Qualifiers		/note="assembly fragment:00105
	1. .258567		fragment_chain:6"
	/organism="Mus musculus"		176847. .195914
	/mol_type="Genomic DNA"		/note="assembly fragment:00636
	/db_xref="taxon:10090"		fragment_chain:7"
	/chromosome="4"		196015. .198745
	/clone="RP23-30108"		/note="assembly fragment:00033
			fragment_chain:7"

FEATURES
source

	<div><div><div><div>protein_id="AAP55137.1"</div><div>db_xref="GI:31433653"</div><div>translation="MEAFWASADLPAFPWGAASLTPPPPPPPHHHHHQOQQOVLPPPA AAPAAAFNQDTLQRLQSIIGSRETWTYIAFWOSSIDVSTGSLGGDGYKGC DDRRQRSSTPAARAEQHRKRVLRSLNSLAGAGAEPDEAVEEBVTPTLWFFVLSMT QSFNGLGLPGCALFAAQTWTATGSLAPCDRAQAYTFLGRTWTVCLPLTFLVLELG SDTVLFTQGDSIPRTRALFNLSAAASWPHPDAAADPSVLWLADAPMDMKDIS AADIISPPPPPHQIQHFENGSTSTLTENPSPVHAPTQSPAAPRQPOQQOQSS QAQQGFFRRLNFGFASNGAAAPFFKPTGTEILNFGNDSSSSRNNPAATANEGLMFS SLTTPAGTSGTQHTTLTAANDAKSNQKSMESATSRASNTNNHPPAATANEGLMFS SAPTTPSTGTCAPAKSDSDLEASREVESRVVAPPEAEKPRKRKPKPANGR EPLNHAERORRKLNORFVALRVVNVSKMDKASLGLDAIYINELRKLTALE TKETLQSOESLKKERDARPPAPSGGGDGCARGCHAVEIAKIIGLSGLMVRQCHKR NHPAARLMTALRELDLDVYHASVSVVKDLMIQOVAVKMASRVYSQDLNAALYTIAE PQTAAR"</div><div><50630. .>50700</div><div>/locus_tag="CHR10v09222002.LRNA-Gly-2"</div><div>50630. .50700</div><div>/locus_tag="CHR10v09222002.LRNA-Gly-2"</div><div>/product="tRNA-Gly"</div><div>complement(<51371. .>65531)</div><div>/locus_tag="OSUNBa0027L23.1"</div><div>/note="contains Pfam profile PF00226 (DnaJ domain)"</div><div>complement(join(<51371. .51598,51735. .51803,52273. .52396, 52749. .52873,53371. .53832,53913. .54256,54617. .54962, 55048. .55497,56316. .56807,56727. .57071,57167. .57390, 57628. .58119,58220. .58612,59157. .59297,59401. .59763, 60050. .62239,62854. .62916,62990. .63091,63228. .63583, 63882. .63969,64073. .64306,65064. .>65531))</div><div>/locus_tag="OSUNBa0027L23.1"</div><div>complement(join(51371. .51598,51735. .51803,52273. .52396, 52749. .52873,53371. .53832,53913. .54256,54617. .54962, 55048. .55497,56316. .56807,56727. .57071,57167. .57390, 57628. .58119,58220. .58612,59157. .59297,59401. .59763, 60050. .62239,62854. .62916,62990. .63091,63228. .63583, 63882. .63969,64073. .64306,65064. .>65531))</div><div>/locus_tag="OSUNBa0027L23.1"</div><div>/codon_start=1</div><div>/product="putative DnaJ domain containing protein"</div><div>/protein_id="AAP55138.1"</div><div>db_xref="GI:31433654"</div><div>translation="WDFASRAASAPPPAPDASSPSSPRASSSSSAASAEPEYLA RYFVKHNRGYRRLICASSGLVTLDPATLVNTSDASVGFDRAPEGNATFTLL TUTDRAFKALRFPSPURAGILTELRLRPVHPVLDPPVLHLRRHTEWAPFLKV TSVGEELHGHGSDRLWCLDFRMDSPAILIGDSYGRKTAGGGFVLCPLYGRKSKA FMAAGSTNTLIISYLTKTANGSVGVSHVDNSQSMATDFIAKANEAVGAATRHG EWSVRLRPAAGHTASIESLSIGIGRGLGHDGVSQRLVLTNTSLVRRPENYEV DVGFTSPVLIOTIRKISITYEHASTFYTPCQINLLIVVSRDEAGHPWAIIVRPLSTV SALVRPAEPQAFPEFNDGCPHVVASTSRDLATVLDVLQNSOCAIPVLPRLTM PGHRDPCGVARVHSHTVDMEAASHIKHLATVAKAVSVSDTVPGAKIRLWRI REFNACIPYGVPIVIEVPEVVMALISLLPATPNQLPADAPLPPLPPSPAAATIMGF VACLRLLTSRVSARHVAFPVAVGRIMGLNRSEGVAAEAGLVAMLIGGGPGDTS MLTDTRGESHATYMAKSVLFQPVVPIVLNRLRPLSPILLSISIVLEAMLCPE HGETTOHATFELLROVAGRLRLFALFAHPAESVRETVMRTIAERDAIAAESMR DAALDKGALLHLNAFFYPAGERDVSQQLVALWADSYQPADLLLSRLPLGLVAYL HTRSDSDSNQDVEPLSRQRKILQQRSRGCKSMAPVEQCPMSNNEGDFEGHTNV GPGADVHORHQAHPYPTPGISIDPSQVPGVFEAFYENNHTQGAQLQDLSHA YLVDNSNGDLANSALHSDFSVPAQVVENTPVGSGLLCNWYGFRAFSLDHNRAALI WNRTRQELREALDQAEHSLDVEKERTDDIPVGSSTVEDASDSETLPRISNVFVSFV SYPLSKEVSGQYLRLLLESGSNVRAODFLRPDPVAFPRALYHFLCDADIGTVD GAVPDLGCDMDCDLGRLDGFGGGGSGVRELCSFAMAIVVEQHNKVIPDPDGTAIH TVLDDTDRLALNHRLLILLKALMDLNSVVCVLVGGCVLADVMTLVHAESERTAI PQSNLIATAFMEPLKEWMIIDKGQVPLEKDAIRLWSKSIDWTIKCWSGMS DKWLRDIRDELWALAVKVPVLTPTSQIGDAALSILHMSASHSDLDADGEIVTPTPR KRILSPCLPVAQMLTCEPSIVEAAASLLKAIIVTRNPKAMIRLYSTGAFYFALAY PGSNLLSIAQLFSATHQAFFGCEBAVSSSLPLAKRSVIGGLLPESLIVLYLSSGP SAFAAMVSDSDTPEIITWTKWRAEHLIRQVLQHLGDPFPQKLAQHCHSLDYVAPWPPV TYPNLKDEMCWCHRYLRNLCDLIRFFNFIVHEVFEQLSLLAMWREELTRRPMDLSEE</div></div></div></div>	<div><div><div><div>1</div><div>AAAAAACGACACAGCGAGCCTCGG</div><div>25</div></div><div> </div><div> </div></div><div><div><div><div>286923</div><div>AAAAAGGACACAGCTAGCACTCAG</div><div>286899</div></div><div> </div><div> </div></div></div><div><div><div><div>RESULT 30</div><div>BX322621</div></div><div>LOCUS</div><div>DEFINITION</div><div>BX322621</div><div>142359 bp</div><div>DNA</div><div>linear</div><div>HTG 09-JUL-2004</div><div>Danio rerio clone RP71-80B24, WORKING DRAFT SEQUENCE, 4 unordered pieces.</div></div><div><div><div><div>ACCESSION</div><div>BX322621</div></div><div>VERSION</div><div>GI:50199216</div></div><div><div><div><div>KEYWORDS</div><div>HTG; HTGS_PHASE1; HTGS_ACTIVEFIN; HTGS_DRAFT; HTGS_FULLTOP.</div></div><div>SOURCE</div><div>Danio rerio (zebrafish)</div></div><div><div><div><div>ORGANISM</div><div>Danio rerio</div></div><div>Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.</div></div></div><div><div><div><div>REFERENCE</div><div>1</div><div>(bases 1 to 142359)</div></div><div>AUTHORS</div><div>Lloyd,D.</div></div><div><div><div><div>TITLE</div><div>Direct Submission</div></div><div>JOURNAL</div><div>Submitted (08-JUN-2004) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk On Jul 10, 2004 this sequence version replaced gi:49658399.</div></div></div><div><div><div><div>COMMENT</div><div>Center: Wellcome Trust Sanger Institute</div></div><div>Center code: SC</div><div>Web site: http://www.sanger.ac.uk</div><div>Contact: zfish-help@sanger.ac.uk</div><div>----- Project Information</div><div>Center project name: bz80B24</div><div>----- Summary Statistics</div><div>Assembly program: XGAP4; version 4.5</div><div>Chemistry: Dye-terminator; 100% of reads</div><div>Consensus quality: 117499 bases at least Q40</div><div>Consensus quality: 117570 bases at least Q30</div><div>Consensus quality: 117601 bases at least Q20</div><div>Insert size: 142059; sum-of-contigs</div><div>Insert size: 125803; 4.8% error; agarose-fp</div><div>Quality coverage: 10.26x in Q20 bases; sum-of-contigs Quality coverage: 11.70x in Q20 bases; agarose-fp</div><div>-----</div><div>* NOTE: This is a 'working draft' sequence. It currently consists of 4 contigs. The true order of the pieces</div><div>* is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as</div><div>* runs of N, but the exact sizes of the gaps are unknown.</div><div>* This record will be updated with the finished sequence</div><div>* as soon as it is available and the accession number will</div><div>* be preserved.</div><div>* 1 9084: contig of 9084 bp in length</div><div>* 9085 9184: gap of 100 bp</div><div>* 9185 81290: contig of 72106 bp in length</div><div>* 81291 81390: gap of 100 bp</div><div>* 81391 105790: contig of 24400 bp in length</div><div>* 105791 105890: gap of 100 bp</div><div>* 105891 142359: contig of 36469 bp in length.</div><div>Location/Qualifiers</div><div>1. .142359</div><div>/organism="Danio rerio"</div><div>/mol_type="genomic DNA"</div><div>/db_xref="taxon:7955"</div><div>/clone_lib="RP71-80B24"</div><div>/clone_lib="RPCI-71"</div><div>1. .9084</div><div>/note="assembly fragment:01685</div><div>fragment_chain:1"</div><div>9185. .81290</div><div>/note="assembly fragment:02193</div><div>fragment_chain:1"</div><div>81391. .105790</div><div>/note="assembly fragment:02272"</div></div></div></div></div><div data-bbox="1455 88 1555 2011" data-label="Page-Footer"><div><div>Query Match</div><div>Best Local Similarity</div><div>Matches</div></div><div><div>74.4%;</div><div>84.0%;</div><div>21; Conservative</div></div><div><div>Score 18.6;</div><div>Pred. No. 5.4e+02;</div><div>4; Indels</div></div><div><div>Length 300029;</div><div>0; Gaps</div><div>0;</div></div></div></div></div></div>
--	--	--

```

misc_feature 105891..142359
/note="assembly_fragment:02115
clone_end:SP6
vector_side:right"

ORIGIN
Query Match 73.6%; Score 18.4; DB 14; Length 142359;
Best Local Similarity 95.0%; Pred. No. 7e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AAAAAACGACACGCGAGCCC 21
||||| ||||| ||||| |||||
Db 26886 AAAAAAGACACGCGAGCCC 26905

RESULT 31
CR936300 158420 bp DNA linear HTG 10-AUG-2005
LOCUS
DEFINITION
DANIO rerio chromosome 22 clone CH211-63J24, WORKING DRAFT
SEQUENCE.
ACCESSION
CR936300
VERSION
HTG; HTGS PHASE2; HTGS ACTIVEFIN; HTGS_DRAFT; HTGS_FULLTOP.
KEYWORDS
DANIO rerio (zebrafish)
SOURCE
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
REFERENCE
1 (bases 1 to 158420)
McGuire S.
Direct Submission
Submitted (09-AUG-2005) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
zfish-help@sanger.ac.uk
Project Information
http://www.sanger.ac.uk/projects/D_rerio/faqs.shtml#dataeight
On Aug 12, 2005 this sequence version replaced gi:71891561.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: zfish-help@sanger.ac.uk
----- Project Information
Center project name: zC63J24
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; 100% of reads
Consensus quality: 158405 bases at least Q40
Consensus quality: 158417 bases at least Q30
Consensus quality: 158419 bases at least Q20
Insert size: 158420; sum-of-contigs
Insert size: 168521; 6.4% error; agarose-fp
Quality coverage: 8.66x in Q20 bases; sum-of-contigs Quality
coverage: 8.20x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
Location/Qualifiers
1..158420
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/chromosome="22"
/clone="CH211-63J24"
/clone_lib="CHORI-211"
1..158420
/note="assembly_fragment:01444"

misc_feature
1..158420
/note="assembly_fragment:01444"

ORIGIN
Query Match 73.6%; Score 18.4; DB 14; Length 158420;
Best Local Similarity 95.0%; Pred. No. 7e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

FEATURES
source
1..158420
/organism="Cucumis sativus"
/mol_type="mRNA"
/db_xref="taxon:3659"
/clone="CRR26-5"
/tissue_type="Cotyledon"
/clone_lib="H.Teramoto"
/dev_stage="Seedling"
<1..>252
/codon_start=3
/product="catalase"
/protein_id="BAA09702.1"
/db_xref="GI:1199478"
/translation="ERVVHARGASAKGFLRVTHDITNLSCADFLRAPGVQTPVIVRFS
TVIHERGSPETLDRPGPAVKFTYREVRLVVYAFICT"

QY 3 AAAAAACGACACGCGAGCCCCTCGG 25
||||| ||||| ||||| |||||
Db 201 AAAAATTCACACGCGAGCCCCTCGG 179

RESULT 33
HSA323695 1181 bp DNA linear PRI 18-JUL-2002
LOCUS
DEFINITION
Homo sapiens genomic sequence surrounding NotI site, clone
NR3-161S.
ACCESSION
AJ323695
VERSION
AJ323695.1
KEYWORDS
Homo sapiens (human)
SOURCE
ORGANISM
Homo sapiens

```

```

QY 2 AAAAAACGACACGCGAGCCC 21
||||| ||||| ||||| |||||
Db 66981 AAAAAAGACACGCGAGCCC 67000

```

```

RESULT 32
CUSCC/c
LOCUS
DEFINITION
Cucumis sativus mRNA for catalase, partial cds, clone CRR26-5'.
ACCESSION
D63386
VERSION
D63386.1
KEYWORDS
GI:1199477
SOURCE
Cucumis sativus (cucumber)
ORGANISM
Cucumis sativus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids I; Cucurbitales; Cucurbitaceae; Cucumis.

```

```

REFERENCE
1 (sites)
Toyama,T., Teramoto,H., Takeba,G. and Tsuji,H.
Cytokinin induces a rapid decrease in the levels of mRNAs for
catalase, 3-hydroxy-3-methylglutaryl CoA reductase, lectin and
other unidentified proteins in etiolated cotyledons of cucumber
Plant Cell Physiol. 36 (7), 1349-1359 (1995)
8564304
2 (bases 1 to 252)
Toyama,T., Teramoto,H., Takeba,G. and Tsuji,H.
Unpublished
3 (bases 1 to 252)
Toyama,T.
Direct Submission
Submitted (10-JUL-1995) Tomoko Toyama, Kyoto University, Faculty of
Science, Department of Botany; Kitashirakawa Oiwakekyo, Sakyo-ku,
Kyoto, Kyoto 606-01, Japan
[E-mail:as186@esakura.kudpc.kyoto-u.ac.jp, Tel:075-753-4147,
Fax:075-701-3262]
Location/Qualifiers
1..252
/organism="Cucumis sativus"
/mol_type="mRNA"
/db_xref="taxon:3659"
/clone="CRR26-5"
/tissue_type="Cotyledon"
/clone_lib="H.Teramoto"
/dev_stage="Seedling"
<1..>252
/codon_start=3
/product="catalase"
/protein_id="BAA09702.1"
/db_xref="GI:1199478"
/translation="ERVVHARGASAKGFLRVTHDITNLSCADFLRAPGVQTPVIVRFS
TVIHERGSPETLDRPGPAVKFTYREVRLVVYAFICT"

```

```

REFERENCE
1 (bases 1 to 252)
Toyama,T., Teramoto,H., Takeba,G. and Tsuji,H.
Unpublished
3 (bases 1 to 252)
Toyama,T.
Direct Submission
Submitted (10-JUL-1995) Tomoko Toyama, Kyoto University, Faculty of
Science, Department of Botany; Kitashirakawa Oiwakekyo, Sakyo-ku,
Kyoto, Kyoto 606-01, Japan
[E-mail:as186@esakura.kudpc.kyoto-u.ac.jp, Tel:075-753-4147,
Fax:075-701-3262]
Location/Qualifiers
1..252
/organism="Cucumis sativus"
/mol_type="mRNA"
/db_xref="taxon:3659"
/clone="CRR26-5"
/tissue_type="Cotyledon"
/clone_lib="H.Teramoto"
/dev_stage="Seedling"
<1..>252
/codon_start=3
/product="catalase"
/protein_id="BAA09702.1"
/db_xref="GI:1199478"
/translation="ERVVHARGASAKGFLRVTHDITNLSCADFLRAPGVQTPVIVRFS
TVIHERGSPETLDRPGPAVKFTYREVRLVVYAFICT"

```

```

FEATURES
source
1..252
/organism="Cucumis sativus"
/mol_type="mRNA"
/db_xref="taxon:3659"
/clone="CRR26-5"
/tissue_type="Cotyledon"
/clone_lib="H.Teramoto"
/dev_stage="Seedling"
<1..>252
/codon_start=3
/product="catalase"
/protein_id="BAA09702.1"
/db_xref="GI:1199478"
/translation="ERVVHARGASAKGFLRVTHDITNLSCADFLRAPGVQTPVIVRFS
TVIHERGSPETLDRPGPAVKFTYREVRLVVYAFICT"

```

```

FEATURES
source
1..252
/organism="Cucumis sativus"
/mol_type="mRNA"
/db_xref="taxon:3659"
/clone="CRR26-5"
/tissue_type="Cotyledon"
/clone_lib="H.Teramoto"
/dev_stage="Seedling"
<1..>252
/codon_start=3
/product="catalase"
/protein_id="BAA09702.1"
/db_xref="GI:1199478"
/translation="ERVVHARGASAKGFLRVTHDITNLSCADFLRAPGVQTPVIVRFS
TVIHERGSPETLDRPGPAVKFTYREVRLVVYAFICT"

```

```

FEATURES
source
1..252
/organism="Cucumis sativus"
/mol_type="mRNA"
/db_xref="taxon:3659"
/clone="CRR26-5"
/tissue_type="Cotyledon"
/clone_lib="H.Teramoto"
/dev_stage="Seedling"
<1..>252
/codon_start=3
/product="catalase"
/protein_id="BAA09702.1"
/db_xref="GI:1199478"
/translation="ERVVHARGASAKGFLRVTHDITNLSCADFLRAPGVQTPVIVRFS
TVIHERGSPETLDRPGPAVKFTYREVRLVVYAFICT"

```

```

FEATURES
source
1..252
/organism="Cucumis sativus"
/mol_type="mRNA"
/db_xref="taxon:3659"
/clone="CRR26-5"
/tissue_type="Cotyledon"
/clone_lib="H.Teramoto"
/dev_stage="Seedling"
<1..>252
/codon_start=3
/product="catalase"
/protein_id="BAA09702.1"
/db_xref="GI:1199478"
/translation="ERVVHARGASAKGFLRVTHDITNLSCADFLRAPGVQTPVIVRFS
TVIHERGSPETLDRPGPAVKFTYREVRLVVYAFICT"

```

```

FEATURES
source
1..252
/organism="Cucumis sativus"
/mol_type="mRNA"
/db_xref="taxon:3659"
/clone="CRR26-5"
/tissue_type="Cotyledon"
/clone_lib="H.Teramoto"
/dev_stage="Seedling"
<1..>252
/codon_start=3
/product="catalase"
/protein_id="BAA09702.1"
/db_xref="GI:1199478"
/translation="ERVVHARGASAKGFLRVTHDITNLSCADFLRAPGVQTPVIVRFS
TVIHERGSPETLDRPGPAVKFTYREVRLVVYAFICT"

```

```

FEATURES
source
1..252
/organism="Cucumis sativus"
/mol_type="mRNA"
/db_xref="taxon:3659"
/clone="CRR26-5"
/tissue_type="Cotyledon"
/clone_lib="H.Teramoto"
/dev_stage="Seedling"
<1..>252
/codon_start=3
/product="catalase"
/protein_id="BAA09702.1"
/db_xref="GI:1199478"
/translation="ERVVHARGASAKGFLRVTHDITNLSCADFLRAPGVQTPVIVRFS
TVIHERGSPETLDRPGPAVKFTYREVRLVVYAFICT"

```

```

FEATURES
source
1..252
/organism="Cucumis sativus"
/mol_type="mRNA"
/db_xref="taxon:3659"
/clone="CRR26-5"
/tissue_type="Cotyledon"
/clone_lib="H.Teramoto"
/dev_stage="Seedling"
<1..>252
/codon_start=3
/product="catalase"
/protein_id="BAA09702.1"
/db_xref="GI:1199478"
/translation="ERVVHARGASAKGFLRVTHDITNLSCADFLRAPGVQTPVIVRFS
TVIHERGSPETLDRPGPAVKFTYREVRLVVYAFICT"

```

```

FEATURES
source
1..252
/organism="Cucumis sativus"
/mol_type="mRNA"
/db_xref="taxon:3659"
/clone="CRR26-5"
/tissue_type="Cotyledon"
/clone_lib="H.Teramoto"
/dev_stage="Seedling"
<1..>252
/codon_start=3
/product="catalase"
/protein_id="BAA09702.1"
/db_xref="GI:1199478"
/translation="ERVVHARGASAKGFLRVTHDITNLSCADFLRAPGVQTPVIVRFS
TVIHERGSPETLDRPGPAVKFTYREVRLVVYAFICT"

```

```

FEATURES
source
1..252
/organism="Cucumis sativus"
/mol_type="mRNA"
/db_xref="taxon:3659"
/clone="CRR26-5"
/tissue_type="Cotyledon"
/clone_lib="H.Teramoto"
/dev_stage="Seedling"
<1..>252
/codon_start=3
/product="catalase"
/protein_id="BAA09702.1"
/db_xref="GI:1199478"
/translation="ERVVHARGASAKGFLRVTHDITNLSCADFLRAPGVQTPVIVRFS
TVIHERGSPETLDRPGPAVKFTYREVRLVVYAFICT"

```

```

FEATURES
source
1..252
/organism="Cucumis sativus"
/mol_type="mRNA"
/db_xref="taxon:3659"
/clone="CRR26-5"
/tissue_type="Cotyledon"
/clone_lib="H.Teramoto"
/dev_stage="Seedling"
<1..>252
/codon_start=3
/product="catalase"
/protein_id="BAA09702.1"
/db_xref="GI:1199478"
/translation="ERVVHARGASAKGFLRVTHDITNLSCADFLRAPGVQTPVIVRFS
TVIHERGSPETLDRPGPAVKFTYREVRLVVYAFICT"

```

```

FEATURES
source
1..252
/organism="Cucumis sativus"
/mol_type="mRNA"
/db_xref="taxon:3659"
/clone="CRR26-5"
/tissue_type="Cotyledon"
/clone_lib="H.Teramoto"
/dev_stage="Seedling"
<1..>252
/codon_start=3
/product="catalase"
/protein_id="BAA09702.1"
/db_xref="GI:1199478"
/translation="ERVVHARGASAKGFLRVTHDITNLSCADFLRAPGVQTPVIVRFS
TVIHERGSPETLDRPGPAVKFTYREVRLVVYAFICT"

```

```

FEATURES
source
1..252
/organism="Cucumis sativus"
/mol_type="mRNA"
/db_xref="taxon:3659"
/clone="CRR26-5"
/tissue_type="Cotyledon"
/clone_lib="H.Teramoto"
/dev_stage="Seedling"
<1..>252
/codon_start=3
/product="catalase"
/protein_id="BAA09702.1"
/db_xref="GI:1199478"
/translation="ERVVHARGASAKGFLRVTHDITNLSCADFLRAPGVQTPVIVRFS
TVIHERGSPETLDRPGPAVKFTYREVRLVVYAFICT"

```



```
/product="putative Protein export membrane protein
SecD/SecE"
/db_xref="GI:24194980"
/protein_id="AAN48596.1"
/translation="MKQFLSRITSDILNPIRSCSILFVLILLSSFWQSKULTVNSNNL
DILLKPNVSVTKQVIMIGWNGFYILSIKFKDEKMTDHLVAFVARKGQPEIAE
KELKAELKQONVSYKEREKAIKASDILNERLLKEKKFVQVYSYRNVSLQKARDI
PLFLTEDLIEVRKEVKIKIDEEVERANPEFKLSNEVRNPDNDILSKYOKLAKRDI
FDEYNISDKMLFLIKPAGSFNIEFNIALDKKIKIIVADLGLDRNGIOVGTGY
RUHDDYTDMAALPIATISFGIALLLFFFRNPLFILILLISLGLSGLIFSPGLIT
IVIGLQSVTSIIASLMLGIDYQIQLYFRFEFTREQNTLSIKDTIYHTGIASF
ISALTTFVAFSEFSEFGIATYGLIILVIMISFNFSPGICQNYNGRDLMDV
SKFLTSVKEQSPFLYRFPKPGELLTVIIVIMISFNFSPGICQNYNGRDLMDV
NLDVNLDEIGDRDIDSQPOVIVVDLSEAVEDYMTVPDPIAGSDVQVSLWN
FVPPGQQRANKILKQSDKPKVASFLRPEQRKYLFPVKYKYNLNVKEYSLSEVPYI
FSSQGTVEVKSKEGHLVFIYPKVALMHGQKLLKFFDAVGEHLHPKRRRLVNTLLYD
SNGHYATNPTDKTQNEKRLIVNSITVSASQFNKGLLDATISFILETFPSFLEQ
ARSHKYISNTAGSLIFANLIKIVQREGVIAFLSTLVLVIVLILFFRGIVPALISLI
PLVGLIFVTLGMAAFRIQLNPMNVLPVPIIGYGIQNGIYIYFRFDHVDIVRAMAM
VGPALIASLTTLVCGWSSLLIADQGLKISIGIVASIGIGSSLIILATLVPVLEIVYR
SRKEEKESKPVGFQDAENSSGNLLFASKYTAENLKTETIENSKIATSKSKSVKKK
TTVSKSIQKKK"
complement(4708..5364)
/locus_tag="LA1398"
complement(4708..5364)
/locus_tag="LA1398"
/codon_start=1
/transl_table=11
/product="putative outer membrane protein"
/protein_id="AAN48597.1"
/db_xref="GI:24194981"
/translation="WNSLRHKLSQIKAPIVKKIILFVLISLIPSLIAQTSTENSTS
PETGNPATTEAPDSEEQIVSTVKKLIGFIKYNKDKRAILHVKQFSNQLKSSGK
IGDSRKEFEAEISFTHRSPIAHKYPDKDINYEKPLKGNATLASSIITWNGSE
RITFSWILMKIEGSWYVTDFLNEGKYASTNRVKSIDPSIKKNGVKYQTIATLKKEAN
"
5548..5688
/locus_tag="LA1399"
5548..5688
/locus_tag="LA1399"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAN48598.1"
/db_xref="GI:24194982"
/translation="MEPLSMNFEFDPIRWFIRDFIKFTFFKSRMSGCCLSNRNDKLL
LY"
7375..9321
/locus_tag="LA1400"
7375..9321
/locus_tag="LA1400"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAN48599.1"
/db_xref="GI:24194983"
/translation="MSYVFIAYNVSVIFQKNGYKMGRIWIVRVSLLVLIGIGFBEYGI
NHTSNASSKDSYSTAQPADQPKDSIKVNVHGGSNYCYSVPFTKGEYIWDYCRD
NTAKARYDFVQISYINNNITWLCITAPETVWKEETWYNLPCTINDPLQRIWIKD
NSFQTANGRYRLKDTWYGVYIRNSGDRYDHTLDSMDDWIETVAPIGNISQTSIAW
DIQWTEGNERFYIRKNGSSKNKNTPLYNPESHGIIAQYDPSGLLNCYSKMTDKDWN
WYWKGSQAPIKQONPAPNWFVFETDKSGATDYKGNVIRVTRVGLNWGVAYTKPS
YLEQDTHFTSFLFVTDKDLDRYTSYNSLTKDQYCPAGNKESLGRKVRKRNLLNP
SDFQTRFVWIPRTISFNQYIARNSISGTPIRRGVCGVCMHYSQYIAELLYHSRGLPSTGG
YFDPALNDRPTISFNQYIARNSISGTPIRRGVCGVCMHYSQYIAELLYHSRGLPSTGG
RSGRLHARTLNSPGSIWGLIIRGNANGSTWGHAVPILRTSQIIVIPNTVLTMS
LNTYIRSLAPMDPNEVINRLENGNTLTTLTIQPIRYIDIPFSLTVSTREDCTGDGDG
RKGSGRYPTSSLINQCSCGRCILQ"
complement(9507..9626)
/locus_tag="LA1401"
complement(9507..9626)
/locus_tag="LA1401"
/codon_start=1
```

```
/transl_table=11
/product="hypothetical protein"
/protein_id="AAN48600.1"
/db_xref="GI:24194984"
/translation="MFQILLKSKNTVYLLFEVNSYNSNSELKINFIILFF"

ORIGIN
Query Match 72.8%; Score 18.2; DB 1; Length 10029;
Best Local Similarity 87.0%; Pred. No. 1e+03;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAACGACACGACGACGCCCTC 23
|||||
DB 5131 AAAAAACGACAAAGCGATCGCTC 5109

RESULT 37
AC142512/c
LOCUS
DEFINITION
AC142512 37857 bp DNA linear HTG 03-APR-2003
Rattus norvegicus clone CH230-unknown, *** SEQUENCING IN PROGRESS
*** 23 unordered pieces.
AC142512
HTG; HTGS PHASE1.
Rattus norvegicus (Norway rat)
SOURCE
ORGANISM
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Rattus; Rattus.
Muzny, D. Marie., Metzker, M. Lee., Abramson, S., Adams, C., Alder, J.,
Allen, C., Allien, H., Alebrooks, S., Amin, A., Anguiano, D.,
Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,
Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G.,
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
Gebregorgis, B., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M.,
Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Liu, J.,
Kowis, C., Kraft, C. L., Lebow, H., Levay, J., Lewis, L., Li, Z., Liu, J.,
Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
Lorensuewa, L., Louisedge, H., Lozado, R. J., Lu, X., Ma, J.,
Maheshwari, M., Mahindartne, M., Mahmood, M., Malloy, K., Mangum, A.,
Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
Mawhney, S., McLeod, M., McNeill, T., Meenen, E., Milosavljevic, A.,
Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K.,
Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D.,
Newton, N., Nguyen, N., Norris, S., Nwaokemele, O., Okunolu, G.,
Olarunpugaon, A., Pal, S., Parks, K., Pasternak, S., Paul, H.,
Peretz, A., Perez, L., Pfankoch, C., Plopper, F., Poindexter, A.,
Popovic, D., Primus, E., Pu, L., Pu, M., Quiroz, J., Rachlin, E.,
Reeves, K., Regier, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y.,
Reuter, M., Richards, S., Riggs, E., Rives, C., Rodkey, T., Rojas, A.,
Rose, M., Rose, R., Ruiz, S. J., Sanders, W., Savary, G., Scherer, S.,
Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A.,
Sisson, I., Sitter, C. D., Smaj, D., Sneed, A., Sodergren, E.,
Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A.,
Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S.,
Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villasa, D.,
Waldron, L., Walker, B., Wang, J., Wang, S., Warren, J.,
```


RESULT 38

AC101757/c

LOCUS

DEFINITION

AC101757

VERSION

KEYWORDS

SOURCE

ORGANISM

AC101757 148875 bp DNA linear ROD 26-MAY-2004
Mus musculus chromosome 8, clone RP24-545I11, complete sequence.

AC101757 HTG. AC101757.12 GI:47679264

Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 148875)

Birren,B., Nusbaum,C. and Lander,E.

Mus musculus chromosome 8, clone RP24-545I11

2 (bases 1 to 148875)

Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,

Anderson,S., Barna,N., Bastien,V., Boguslavskiy,L., Boukhgalter,B.,

Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,

Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,A.,

Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,

Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,

Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,

Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,

Jones,C., Kamat,A., Karatas,A., Kells,C., Lakoczek,K.,

Lamazares,R., Lander,T., Lehoczy,J., Levine,R., Liu,G.,

MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,

McCarthy,M., McEwan,P., McKernan,K., McPeeters,R., Meldrim,J.,

Meneus,L., Mihova,T., Mienga,V., Murphy,T., Naylor,J., Nguyen,C.,

Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,

Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,

Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,

Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,

Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,

Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,

Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,

Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,

Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission

Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

3 (bases 1 to 148875)

Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,

Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V.,

Bloom,T., Boguslavskiy,L., Boukhgalter,B., Camarata,J., Chang,J.,

Choepel,Y., Collymore,A., Cooke,A., Cooke,P., Corum,B.,

DeArellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L.,

Erickson,J., Faro,S., Ferreira,P., FitzGerald,M., Gage,D.,

Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N.,

Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I.,

Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Lander,T.,

Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbitt,R.,

MacLean,C., Macdonald,P., Major,J., Manning,J., Matthews,C.,

McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,

Murphy,T., Naylor,J., Nguyen,C., Nguyen,T., Nicol,R., Norbu,C.,

O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,

Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C.,

Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupback,R.,

Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,

Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J.,

Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R.,

Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L.,

Zimmer,A. and Zody,M.

Direct Submission

Submitted (21-APR-2004) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

4 (bases 1 to 148875)

Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,

Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V.,

Bloom,T., Boguslavskiy,L., Boukhgalter,B., Camarata,J., Chang,J.,

Choepel,Y., Collymore,A., Cooke,A., Cooke,P., Corum,B.,

DeArellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L.,

Erickson,J., Faro,S., Ferreira,P., FitzGerald,M., Gage,D.,

Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N.,

Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I.,

DeArellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L.,
Erickson,J., Faro,S., Ferreira,P., FitzGerald,M., Gage,D.,
Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N.,
Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I.,
Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Lander,T.,
Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbitt,R.,
MacLean,C., Macdonald,P., Major,J., Manning,J., Matthews,C.,
McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
Murphy,T., Naylor,J., Nguyen,C., Nguyen,T., Nicol,R., Norbu,C.,
O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C.,
Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R.,
Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L.,
Zimmer,A. and Zody,M.

Direct Submission

Submitted (26-MAY-2004) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

On May 26, 2004 this sequence version replaced gi:46195628.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/MIT Center for Genome Research

Center code: WIGR

Web site: http://www-seq.wi.mit.edu

Contact: sequence_submissions@broad.mit.edu

----- Project Information

Center project name: Li7350

Center clone name: 545_I_11

FEATURES

source

Location/Qualifiers

1..148875

/organism="Mus musculus"

/mol_type="genomic DNA"

/db_xref="taxon:10090"

/chromosome="8"

/map="8"

/clone="RP24-545I11"

/clone_lib="RPCI-24 Male Mouse BAC"

1..4

/note="clone boundary"

clone end:SP6

site:WBI

3564..3684

/rpt_family="PB1D9"

3685..3823

/rpt_family="GA-rich"

3826..4005

/rpt_family="(GGAA)n"

6548..6568

/rpt_family="(T)n"

7260..7305

/rpt_family="(CA)n"

7920..7940

/rpt_family="AT-rich"

8148..8211

/rpt_family="(TCC)n"

complement(8172..8228)

/note="single clone coverage"

8212..8258

/rpt_family="(TC)n"

8259..8290

/rpt_family="(CA)n"

8370..8394

/rpt_family="AT-rich"

complement(9340..9464)

/rpt_family="PB1D7"

complement(9690..9823)

/rpt_family="Bif"

13451..13471

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

```

repeat_region      /rpt_family="AT_rich"
14717..14755
/rpt_family="(TGG)n"
repeat_region      15197..15223
/rpt_family="(CACA)n"
repeat_region      15448..15543
/rpt_family="(CCG)n"
repeat_region      15978..16038
/rpt_family="C-rich"
repeat_region      16987..17058
/rpt_family="(CA)n"
repeat_region      20233..20256
/rpt_family="AT_rich"
repeat_region      20452..20480
/rpt_family="(CCCCCA)n"
repeat_region      20551..20602
/rpt_family="A-rich"
repeat_region      21088..21151
/rpt_family="(GA)n"
repeat_region      22496..22517
/rpt_family="AT_rich"
repeat_region      23614..23690
/rpt_family="(TC)n"
repeat_region      complement(23785..23950)
/rpt_family="RSINE1"
repeat_region      24555..24713
/rpt_family="(TTATA)n"
repeat_region      complement(24923..25029)
/rpt_family="B3"
repeat_region      complement(26778..26965)
/rpt_family="MLT1A1"
repeat_region      complement(27142..27205)
/rpt_family="ID_B1"
repeat_region      complement(27227..27412)
/rpt_family="B2_Mm2"
repeat_region      27610..27652
/rpt_family="(TTG)n"
repeat_region      27861..27794
/rpt_family="PB1D9"
repeat_region      complement(29156..29497)
/rpt_family="Lx5"
repeat_region      29697..29804
/rpt_family="CT-rich"
repeat_region      complement(29952..30222)
/rpt_family="ORR1C"
repeat_region      30261..30404
/rpt_family="B1_MM"
repeat_region      30405..30427
/rpt_family="(CAAAA)n"
repeat_region      30689..30728
/rpt_family="(TTTTG)n"
repeat_region      complement(30969..31106)
/rpt_family="single clone coverage"
repeat_region      31090..31144
/rpt_family="(TG)n"
repeat_region      31668..31697
/rpt_family="(TG)n"

Query Match      72.8%; Score 18.2; DB 9; Length 148875;
Best Local Similarity 87.0%; Pred. No. 8.7e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy      2 AAAAAACGACACGAGCGACCTCG 24
| | | | | | | | | | | | | | | | | |
Db      117549 AGAAACGACACGAGCGACCTGG 117527

RESULT 39
AC137672
LOCUS      AC137672      150313 bp      DNA      linear      PRI 30-JAN-2003
DEFINITION Homo sapiens chromosome 3 clone RP11-72112, complete sequence.
ACCESSION AC137672
VERSION AC137672.2 GI:28144431

```

HTG.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
1 (bases 1 to 150313)
Kaul, R.K., Olson, M.V., Zhou, Y., James, R.A., Rouse, G., Wu, Z.,
Saenphimmachak, C., Buckley, D., Kibukawa, M., Raymond, C. and
Haugen, E.D.
Direct Submission
Unpublished
2 (bases 1 to 150313)
Kaul, R.K., Olson, M.V., Raymond, C. and Haugen, E.D.
Direct Submission
Submitted (27-NOV-2002) Genome Center, University of Washington,
Box 352145, Seattle, WA 98195, USA
3 (bases 1 to 150313)
Kaul, R.K., Olson, M.V., Zhou, Y., James, R.A., Rouse, G., Wu, Z.,
Saenphimmachak, C., Buckley, D., Kibukawa, M., Raymond, C. and
Haugen, E.D.
Direct Submission
Submitted (30-JAN-2003) Genome Center, University of Washington,
Box 352145, Seattle, WA 98195, USA
On Jan 30, 2003 this sequence version replaced gi:25704977.
----- Genome Center
Center: University of Washington Genome Center
Center Code: UWGC
Web site: http://www.genome.washington.edu
Contact: uwgchgs@u.washington.edu
----- Project Information
Center project name: chr-3
Center clone name: RP11-72112 (bc0748)
----- Summary Statistics
Sequencing vector: plasmid; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 150288 bases at least Q40
Consensus quality: 150310 bases at least Q30
Consensus quality: 150313 bases at least Q20
Insert size: 150313; sum-of-contigs
Quality coverage: 7.8x in Q20 bases; sum-of-contigs

Overlapping Sequences:
5': RP11-14762 (UWGC:bc0804) AC132807, 66478-bp overlap
3': RP11-554B20 (UWGC:bc0471) AC099543, 59734-bp overlap

Sequence Quality Assessment:
This entry has been annotated with sequence quality
estimates computed by the Phrap assembly program.
All manually edited bases have been reduced to quality zero.
Quality levels above 40 are expected to have less than
1 error in 10,000 bp.
Base-by-base quality values are not generally visible from the
GenBank flat file format but are available as part
of this entry's ASN.1 file.

This sequence was finished as follows unless otherwise noted:
all regions were either double-stranded or sequenced with an
alternate chemistry or covered by high quality data (i.e., Phred
quality >= 30); an attempt was made to resolve all sequencing
problems, such as compressions and repeats; all regions were
covered by at least one plasmid subclone or more than one M13
subclone; and the assembly was confirmed by restriction digest.

Sequence Validation:
This sequence has been validated by Multiple Complete Digest
fingerprinting. Comparison of the experimentally derived digest
fragments with sequence-predicted fragments is given below.
The electronically-digested sequence consists of both insert and
vector, in order to accurately represent the entire circular BAC.
Small fragments below a variable cutoff (approximately 400-800 bp)
are not resolved in the fingerprint and hence do not appear

KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
AUTHORS
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

in the table. There are no significant remaining discrepancies between the experimental and predicted values. Uniquely ordered fragments are separated by dashed lines.

BgIII				EcoRI				HindIII			
SeqDerMap	FngPrnt	SeqDerMap	FngPrnt	SeqDerMap	FngPrnt	SeqDerMap	FngPrnt	SeqDerMap	FngPrnt	SeqDerMap	FngPrnt
3570	3516	8696	8685	882	903	2282	2349	5534	5485	228	<800
2067	2056	6	<800	6382	6532	2152	2252	1115	1112	4200	4186
5793	5871	2008	1921	512	<800	118	<800	3701	3692	1944	1889
315	<800	111	<800	449	<800	962	971	11627	11463	906	903
3082	3132	502	<800	1879	1889	2878	2885	5970	5989	259	<800
3061	3024	2208	2216	3776	3777	2801	2787	2973	2967	3235	3220
1275	1223	3870	3784	6154	6139	2484	2518	481	<800	9354	9177
2532	2518	2566	2584	690	<800	18	<800	189	<800	625	<800
7238	7231	100	<800	6	<800	39	<800	5947	5989	470	<800
4907	4698	214	<800	2845	2888	749	767	2556	2584	211	<800
325	<800	958	974	169	<800	2370	2518	1056	1051	6170	6139
3292	3365	8629	8685	2853	2888	1603	1561	601	<800	2225	2219
745	<800	2703	2584	1485	1434	1870	1838	4672	4591	4526	4476
1624	1561	3028	2967	1278	1255	1534	1561	1069	1051	910	903
2113	2125	1125	1112	2124	2098	2337	2349	1485	1461	1902	1889
591	<800	2174	2216	524	<800	2111	2125	5643	5674	886	903
8685	8655	5329	5296	533	<800	6710	6732	5112	5021	591	<800
3668	3631	2974	2967	3371	3323	4463	4385	---	---	2451	2467
1840	1838	5981	5989	1889	1889	4807	4698	---	---	1442	1434
660	<800	8695	8685	459	<800	---	---	---	---	---	---
8518	8655	1757	1743	261	<800	3136	3132	---	---	12899	12900
1934	1838	2304	2414	2861	2888	---	---	---	---	4605	4623
5920	5871	1045	1051	206	<800	3337	3365	---	---	2078	2098
3447	3365	8191	8148	574	<800	5507	5466	---	---	320	<800
1866	1838	533	<800	1386	1327	1152	1124	---	---	509	<800
3570	3516	2471	2584	508	<800	---	---	---	---	1588	1576
3674	3631	372	<800	862	903	---	---	---	---	---	---
4060	4023	1002	974	1866	1889	---	---	---	---	---	---
798	767	6502	6455	129	<800	---	---	---	---	---	---
3294	3246	6486	6455	8466	8529	---	---	---	---	---	---
4436	4385	3244	3200	5025	4978	---	---	---	---	---	---
4692	4698	3497	3522	2053	2098	---	---	---	---	---	---

Query Match 72.8%; Score 18.2; DB 8; Length 150313;
Best Local Similarity 87.0%; Pred. No. 8.7e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 AAAACGACACGAGCGCTCGG 25
Db 96576 AAAACGACACGAGCGCTTGG 96598

RESULT 40
AC099543/c

LOCUS AC099543 156107 bp DNA linear PRI 19-APR-2002
DEFINITION Homo sapiens chromosome 3 clone RP11-554B20, complete sequence.

ACCESSION AC099543 AC055718

VERSION AC099543.2 GI:20198581

KEYWORDS HTG.

SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE 1 (bases 1 to 156107)

2313	2280	3937	3851	485	<800
4071	4108	11994	11832	1109	1128
2428	2471	76	<800	3204	3187
456	<800	1050	1058	3927	3843
3917	3953	853	822	6093	6087
630	<800	631	<800	8295	8286
523	<800	86	<800	609	<800
2335	2280	795	<800	1272	1248
1502	1503	2329	2288	15004	14888
5632	5643	481	<800	7351	7265
1585	1503	6885	6690	2421	2373
2313	2280	176	<800	765	<800
5031	4846	599	<800	3021	3058
855	837	2845	2894	1535	1530
6115	6065	7642	7642	1066	1128
1993	2052	981	1058	7161	7265
4182	4108	7874	7642	2429	2373
2183	2160	460	<800	44	<800
2250	2280	1455	1430	4482	4409
374	<800	-----	-----	1444	1530
2177	2160	-----	-----	-----	-----
6776	6632	-----	-----	-----	-----
Query Match	72.8%;	Score 18.2;	DB 8;	Length 156107;	
Best Local Similarity	87.0%;	Pred. No. 8.7e+02;			
Matches	20;	Conservative 0;	Mismatches 3;	Indels 0;	Gaps 0;
QY	3	AAACGACACAGCGAGCCTCGG 25			
Db	150111	AAACGACACAGCGAGCCTTGG 150089			
RESULT 41					
AC083820/c					
LOCUS					
DEFINITION		190122 bp	DNA	linear	HTG 09-OCT-2002
		Rattus norvegicus clone RP32-290H10, ***	SEQUENCING IN PROGRESS		
ACCESSION					
VERSION		AC083820			
KEYWORDS		HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.			
SOURCE		Rattus norvegicus (Norway rat)			
ORGANISM					
REFERENCE					
AUTHORS					
1		(bases 1 to 190122)			
Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-osman, F.R., Allen, C.,					
Alsbrooks, S.L., Amaratunga, H.C., Are, J.R., Ayele, M., Banks, T.,					
Barbaria, J., Benton, J., Bimage, K., Blankenburg, K., Bonnin, D.,					
Bouck, J., Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P.,					
Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C.,					
Carton, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,					

Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C.,					
Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,					
Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A.,					
Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H.,					
Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,					
Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,					
Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,					
Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,					
Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,					
Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J.,					
Hernandez, O., Hodgson, A., Hogues, M., Holloway, C., Hollins, B.,					
Honsi, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E.,					
Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,					
Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,					
Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,					
Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Loulseged, H.,					
Lorazo, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,					
Mathewswari, M., Mapua, P., Martin, R., Martindale, A., Martinez, B.,					
Massey, E., Mahoney, E., McLeod, M.P., Meador, M., Mei, G., Metzker, M.,					
Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,					
Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,					
Nguyen, N., Nickerson, E., Nwokenkwo, S., Oguh, M., Okwuonu, G.,					
Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peary, J., Perez, L.,					
Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y.,					
Rives, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savary, G.,					
Scherer, S., Scott, G., Shen, H., Shoohtari, N., Sisson, I.,					
Sodergren, E., Sonaike, T., Sparks, A., Stanley, H., Stone, H.,					
Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H.,					
Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,					
Usmani, K., Vasquez, L., Vera, V., Villalob, D., Vinson, R., Wang, Q.,					
Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,					
Williams, G., Williamson, A., Wleczyk, R., Wooden, S., Worley, K.,					
Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,					
Weinstock, G. and Gibbs, R.					
Direct Submission					
Unpublished					
2 (bases 1 to 190122)					
Worley, K.C.					
Direct Submission					
Submitted (01-OCT-2000)					
of Molecular and Human Genetics, Baylor College of Medicine, One					
Baylor Plaza, Houston, TX 77030, USA					
3 (bases 1 to 190122)					
Worley, K.C.					
Direct Submission					
Submitted (09-OCT-2002)					
of Molecular and Human Genetics, Baylor College of Medicine, One					
Baylor Plaza, Houston, TX 77030, USA					
The sequence in this assembly is a combination of BAC based reads					
and whole genome shotgun sequencing reads assembled using Atlas					
(http://www.hgsc.bcm.tmc.edu/projects/rat/). As a result, the					
sequence may extend beyond the ends of the clone and there may be					
contigs that consist entirely of whole genome shotgun sequence					
reads. Both end sequences and whole genome shotgun sequence only					
contigs will be indicated in the feature table.					
----- Genom Center					
Center: Baylor College of Medicine					
Center code: BCM					
Web site: http://www.hgsc.bcm.tmc.edu/					
Contact: hgsc-help@bcm.tmc.edu					
----- Project Information					
Center project name: TUAH					
Center clone name: RP32-290H10					
----- Summary Statistics					
Assembly program: Phrap; version 0.990329					
Consensus quality: 165695 bases at least Q40					
Consensus quality: 167639 bases at least Q30					
Consensus quality: 168696 bases at least Q20					
Estimated insert size: 203777; sum-of-contigs estimation					
Quality coverage: 4x in Q20 bases; sum-of-contigs estimation					

* NOTE: Estimated insert size may differ from sequence length					

* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 10 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

* 1 12928: contig of 12928 bp in length
 * 12929 13028: gap of unknown length
 * 12930 159863: contig of 146835 bp in length
 * 159864 171348: gap of unknown length
 * 171349 171448: contig of 11385 bp in length
 * 171449 181129: contig of 9681 bp in length
 * 181130 181229: gap of unknown length
 * 181230 182307: contig of 1077 bp in length
 * 182308 182406: gap of unknown length
 * 182407 183773: contig of 1367 bp in length
 * 183774 185328: gap of unknown length
 * 185329 185428: contig of 1455 bp in length
 * 185429 186735: contig of 1307 bp in length
 * 186736 186835: gap of unknown length
 * 186836 186836: contig of 1801 bp in length
 * 186837 188736: gap of unknown length
 * 188737 190122: contig of 1386 bp in length.

FEATURES

source	Location/Qualifiers
1.	190122
/organism="Rattus norvegicus"	
/mol_type="genomic DNA"	
/db_xref="taxon:10116"	
/clone="RP32-290H10"	
12929..13028	
/estimated_length=unknown	
13029..14076	
/notes="wgs contig"	
159864..159963	
/estimated_length=unknown	
171349..171448	
/estimated_length=unknown	
181130..181229	
/estimated_length=unknown	
182307..182406	
/estimated_length=unknown	
183774..183873	
/estimated_length=unknown	
185329..185428	
/estimated_length=unknown	
186736..186835	
/estimated_length=unknown	
188637..188736	
/estimated_length=unknown	

ORIGIN

Query Match 72.8%; Score 18.2; DB 14; Length 190122;
 Best Local Similarity 87.0%; Pred. No. 8.6e+02;
 Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AAAAAGCAGACAGCGAGCCCTC 23

Db 9832 AAAAAGCAGACAGCGAGCCCTC 9810

RESULT 42

AC080156 218222 bp DNA linear HTG 09-OCT-2002
 LOCUS Rattus norvegicus clone RP32-379B20, *** SEQUENCING IN PROGRESS
 DEFINITION *** 8 unordered pieces.
 AC080156
 AC080156.9 GI:23269308
 VERSION HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
 KEYWORDS

SOURCE ORGANISM

Rattus norvegicus (Norway rat)

Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muridea; Muridae; Murinae; Rattus.

REFERENCE AUTHORS

1 (bases 1 to 212822)
 Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-osman, F.R., Allen, C.,
 Alsbrooks, S.L., Amaratunga, H.C., Are, J.R., Ayale, M., Banks, T.,
 Barbarta, J., Benton, J., Binnage, K., Blankenburg, K., Bonnin, D.,
 Bouck, J., Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P.,
 Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C.,
 Carron, T.P., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,
 Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C.,
 Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,
 Davila, M.L., Davis, C., Davy-Carrroll, L., Dederich, D.A.,
 Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H.,
 Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,
 Earhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,
 Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,
 Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,
 Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,
 Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J.,
 Hernandez, O., Hodgson, A., Hogue, M., Holloway, C., Hollins, B.,
 Homei, F., Howard, S., Huber, J., Hulyk, S., Humet, J., Joudan, S.,
 Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudan, S.,
 Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,
 Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,
 Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Louleaged, H.,
 Lozada, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,
 Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,
 Massey, E., Mawhiney, E., McLeod, M.P., Meador, M., Mei, G., Metzker, M.,
 Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,
 Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,
 Nguyen, N., Nickerson, E., Nwokenkwo, S., Ogih, M., Okwuonu, G.,
 Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,
 Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y.,
 Rives, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savary, G.,
 Scherer, S., Scott, G., Shen, H., Shoohtari, N., Sisson, I.,
 Sodergren, E., Sonaik, T., Sparks, A., Stanley, H., Stone, H.,
 Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H.,
 Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,
 Usmani, K., Vasquez, L., Vera, V., Villalobos, D., Vinson, R., Wang, Q.,
 Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,
 Williams, G., Williamson, A., Wleczky, R., Wooden, S., Worley, K.,
 Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,
 Weinstein, G., and Gibbs, R.

Direct Submission

Unpublished

2 (bases 1 to 212822)

Worley, K.C.

Direct Submission

Submitted (27-SEP-2000) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 212822)

Worley, K.C.

Direct Submission

Submitted (09-OCT-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA

On Sep 23, 2002 this sequence version replaced gi:21953772.

The sequence in this assembly is a combination of BAC based reads
 and whole genome shotgun sequencing reads assembled using Aclis
 (http://www.hgsc.bcm.tmc.edu/projects/rat/). As a result, the
 sequence may extend beyond the ends of the clone and there may be
 contigs that consist entirely of whole genome shotgun sequence
 reads. Both end sequences and whole genome shotgun sequence only
 contigs will be indicated in the feature table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: TUGX
Center clone name: CH230-1L14
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 197323 bases at least Q40
Consensus quality: 201468 bases at least Q30
Consensus quality: 204719 bases at least Q20
Estimated insert size: 219777; sum-of-contigs estimation
Quality coverage: 4x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 245806: contig of 245806 bp in length
* 245807 245906: gap of unknown length
* 245907 247137: contig of 1231 bp in length.

FEATURES

source

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

gap

ORIGIN

Query Match 72.8%; Score 18.2; DB 14; Length 247137;
Best Local Similarity 87.0%; Pred. No. 8.5e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAACGACACAGCGCCCTC 23

Db 182443 AAAAAAGGCCACAGCGGCCCC 182421

RESULT 45

AC096330/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

AC096330
Rattus norvegicus clone CH230-82B5, WORKING DRAFT SEQUENCE, 2
unordered pieces.
AC096330
AC096330.7 GI:30521925
HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
Rattus norvegicus (Norway rat)
Rattus norvegicus

AC098268
AC098268.4 GI:22855777
HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
Rattus norvegicus (Norway rat)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Rattus.

1 (bases 1 to 247137)
Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-oshan, F.R., Allen, C.,
Alzbrooks, S.L., Amarutunge, H.C., Are, J.R., Ayale, M., Banks, T.,
Barbarta, J., Benton, J., Bina, K., Blankenbush, K., Bonnin, D.,
Bouck, J., Bowie, S., Brieva, M., Brown, E., Brown, N.P.,
Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C.,
Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,
Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C.,
Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,
Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A.,
Delaney, K.R., Delgado, O., Denny, A.L., Ding, Y., Dinh, H.H.,
Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,
Earhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,
Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,
Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,
Gortell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,
Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J.,
Hernandez, O., Hodgson, A., Hogue, M., Holloway, C., Hollins, B.,
Homs, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.B.,
Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,
Karleson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,
Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,
Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Loulseghe, H.,
Lozado, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,
Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, B.,
Massey, E., Mawhney, E., McLeod, M.P., Meador, M., Mei, G., Metzker, M.,
Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,
Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,
Nguyen, N., Nickerson, E., Nwokenko, S., Oguh, M., Okunolu, G.,
Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,
Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y.,
Rivers, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savery, G.,
Scherer, S., Scott, G., Shen, H., Shooshitari, N., Sisson, I.,
Sodergren, E., Sonaite, T., Sparks, A., Stanley, H., Stone, H.,
Sutton, A., Svatek, A., Taber, P., Tamerisa, A., Tamerisa, K., Tang, H.,
Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,
Usmani, K., Vasquez, L., Vera, V., Villalob, D., Vinson, R., Wang, Q.,
Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,
Williams, G., Williamson, A., Wleczek, R., Wooden, S., Worley, K.,
Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,
Weinstock, G. and Gibbs, R.

Direct Submission
Unpublished
2 (bases 1 to 247137)
Worley, K.C.

Direct Submission
Submitted (23-OCT-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 247137)
Worley, K.C.

Direct Submission
Submitted (08-OCT-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

On Sep 14, 2002 this sequence version replaced gi:21953869.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE AUTHORS

1 (bases 1 to 249741)
Mullen,D.Marie., Metzker,M.Lee., Abramson,S., Adams,C., Alder,J., Allen,C., Allen,H., Alsbrooks,S., Amin,A., Angulano,D., Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H., Bandwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F., Biewalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M., Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E., Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Chen,Z., Chu,J., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Cline,A., D'Souza,L., Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Amico,L., Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D., Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K., Draper,H., Dugan-Rocha,S., Dunn,A., Dutbin,K., Duval,B., Eaves,K., Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G., Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P., Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M., Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W., Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K., Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J., Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M., Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A., Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A., Karpachy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C., Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J., Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J., Lorenschuhwa,L., Loulseghe,H., Lozado,R.J., Lu,X., Ma,J., Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A., Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E., Mawhney,S., McLeod,M.P., McNeill,T.Z., Meenen,E., Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S., Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L., Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Nwaokeleneh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K., Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkuch,C., Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L., L., Puazo,M., Quiroz,J., Rachin,E., Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F., Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J., Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H., Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajic,D., Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J., Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K., Valas,R., Vera,V., Villanasa,D., Waldron,L., Walker,B., Wang,J., Wang,O., Wang,S., Warren,J., Warren,R., Wei,X., White,F., Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O., Weinstock,G. and Gibbs,R.A.

TITLE JOURNAL

Unpublished

REFERENCE AUTHORS

2 (bases 1 to 249741)
Worley,K.C.
Direct Submission
Submitted (17-SEP-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

TITLE JOURNAL

3 (bases 1 to 249741)

REFERENCE AUTHORS

Rat Genome Sequencing Consortium.
Direct Submission
Submitted (10-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

TITLE JOURNAL

COMMENT

On May 10, 2003 this sequence version replaced gi:24818522. The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold,

individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GEYS
Center clone name: CH230-82B5
----- Summary Statistics
Assembly program: Atlas 3.0;
Consensus quality: 242419 bases at least Q40
Consensus quality: 243103 bases at least Q30
Consensus quality: 243892 bases at least Q20
Estimated insert size: 258836; sum-of-contigs estimation
Quality coverage: 9x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)
* NOTE: This sequence may represent more than one clone.
* NOTE: This is a 'working draft' sequence. It currently consists of 2 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

* 1 248130: contig of 248130 bp in length
* 248131 248230: gap of unknown length
* 248231 249741: contig of 1511 bp in length.

FEATURES source

1..249741
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-82B5"

misc_feature

1..1700
/note="wgs_end_extension"
clone_end:T7
5115..6713
/note="wgs_end_extension"

misc_feature

clone_end:T7
7411..8207
/note="clone_boundary"
clone_end:T7

misc_feature

end sequence:BH356129"
28181..246883
/note="clone boundary"
clone_end:Sp6

misc_feature

site:EcoRI
end sequence:BH356131"
248131..248230
/estimated_length=unknown

gap

ORIGIN

Query Match 72.8%; Score 18.2; DB 14; Length 249741;
Best Local Similarity 87.0%; Pred. No. 8.5e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AAAAAAGCAGCAGCGAGCCCTC 23

Db 185990 AAAAAAGCAGCAGCGAGCCCTC 185968

RESULT 46

```
AC080155/c
LOCUS       271995 bp      DNA      linear      HTG 08-OCT-2002
DEFINITION  Rattus norvegicus clone RP32-324M6, *** SEQUENCING IN PROGRESS ***,
12 unordered pieces.
AC080155
AC080155.9  GI:23269316
VERSION     HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
KEYWORDS    Rattus norvegicus (Norway rat)
SOURCE      Rattus norvegicus
ORGANISM    Rattus norvegicus
REFERENCE   1 (bases 1 to 271995)
AUTHORS     Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-oman,F.R., Allen,C.,
Albrooks,S.L., Amaratunge,H.C., Are,J.R., Ayelle,M., Banks,I.,
Barbata,J., Benton,J., Bimage,K., Blakenburg,K., Bonnin,D.,
Bouck,J., Burch,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carroll,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,W.D., Bathorne,S.R., David,R.,
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Earhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,
Homs,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
Karleson,E., Kelly,S., Khan,J., King,L., Korvah,J., Kovar,C.,
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulsegad,H.,
Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G.,
Ogareny,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savary,G.,
Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sisson,I.,
Sodergren,E., Sonaik,T., Sparks,A., Stanley,H., Stone,H.,
Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
Usmani,K., Vaquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K.,
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 271995)
Worley,K.C.
Direct Submission
Submitted (27-SEP-2000) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 271995)
Worley,K.C.
Direct Submission
Submitted (08-OCT-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Sep 23, 2002 this sequence version replaced gi:22164800.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: TUA4
Center clone name: RP32-324M6
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 224636 bases at least Q40
Consensus quality: 228004 bases at least Q30
Consensus quality: 230671 bases at least Q20
Estimated insert size: 262923; sum-of-contigs estimation
Quality coverage: 3x in Q20 bases; sum-of-contigs estimation
table.
-----
* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)
* NOTE: This sequence may represent more than one clone.
* NOTE: This is a 'working draft' sequence. It currently
consists of 12 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.
* 1 3147: contig of 3147 bp in length
* 3148 3247: gap of unknown length
* 3248 208352: contig of 205105 bp in length
* 208353 208452: gap of unknown length
* 208453 219789: contig of 11337 bp in length
* 219790 219889: gap of unknown length
* 219890 231262: contig of 11373 bp in length
* 231263 231363: gap of unknown length
* 231363 261189: contig of 29827 bp in length
* 261190 261289: gap of unknown length
* 261290 262472: contig of 1183 bp in length
* 262473 262572: gap of unknown length
* 262573 263781: contig of 1209 bp in length
* 263782 263881: gap of unknown length
* 263882 264967: contig of 1086 bp in length
* 264968 265067: gap of unknown length
* 265068 266817: gap of unknown length
* 266818 266917: gap of unknown length
* 266918 268484: contig of 1567 bp in length
* 268485 268584: gap of unknown length
* 268585 270311: contig of 1727 bp in length
* 270312 270411: gap of unknown length
* 270412 271995: contig of 1584 bp in length.
FEATURES
source
1..271995
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="RP32-324M6"
misc_feature
1..1252
/note="wgs_contig"
3148..3247
/estimated_length=unknown
180275..182104
/note="wgs_contig"
182155..184160
/note="wgs_contig"
188211..189659
/note="wgs_contig"
```

```

gap      208353..208452
/misc_feature
gap      208453..210185
/notes="wgs contig"
219790..219889
/estimated_length=unknown
/misc_feature
219890..221345
/notes="wgs contig"
231263..231362
/estimated_length=unknown
misc_feature
236650..238740
/notes="wgs contig"
246937..249470
/notes="wgs contig"
261190..261289
/estimated_length=unknown
gap      262473..262572
/estimated_length=unknown
gap      263782..263881
/estimated_length=unknown
gap      264968..265067
/estimated_length=unknown
gap      266818..266917
/estimated_length=unknown
gap      268485..268584
/estimated_length=unknown
gap      270312..270411
/estimated_length=unknown

```

ORIGIN

```

Query Match      72.8%; Score 18.2; DB 14; Length 271995;
Best Local Similarity 87.08; Pred. No. 8.5e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

Qy 1 AAAAAACGACACGCGAGCCCTC 23

Db 145533 AAAAAAGCCACGCGAGCCCC 145511

```

RESULT 47
AE017296      301258 bp      DNA      linear      BCT 22-MAR-2004
LOCUS
DEFINITION
Leptospira interrogans serovar Copenhageni str. Fioacruz Li-130,
Chromosome I, section 10 of 14 of the complete sequence.
ACCESSION
AE017296
VERSION
AE017296.1 GI:45601342
KEYWORDS
SOURCE
Leptospira interrogans serovar Copenhageni str. Fioacruz Li-130
ORGANISM
Leptospira interrogans serovar Copenhageni str. Fioacruz Li-130
Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.
1 (bases 1 to 301258)
Nascimento,A.L., Ko,A.I., Martins,E.A., Monteiro-Vitorello,C.B.,
Ho,P.L., Haake,D.A., Verjovski-Almeida,S., Hartskeerl,R.A.,
Marques,M.V., Oliveira,M.C., Menck,C.F., Leite,L.C., Carrier,H.,
Coutinho,L.L., Degraeve,W.M., Dellagostin,O.A., El-Dorri,H.,
Ferreto,E.S., Ferro,M.I., Furlan,L.R., Gamberini,M., Gigliotti,E.A.,
Goes-Neto,A., Goldman,G.H., Goldman,M.H., Harakava,R.,
Jeronimo,S.M., Junqueira-De-Azevedo,I.L., Kimura,E.T.,
Kuramae,E.E., Lemos,E.G., Lemos,M.V., Marino,C.L., Nunes,L.R., De
Oliveira,R.C., Pereira,G.G., Reis,M.S., Schriefer,A.,
Siqueira,W.J., Sommer,P., Tsai,S.M., Simpson,A.J., Ferro,J.A.,
Camargo,L.E., Kitajima,J.P., Setubal,J.C. and Van Sluys,M.A.
Comparative Genomics of Two Leptospira interrogans Serovars Reveals
Novel Insights into Physiology and Pathogenesis
J. Bacteriol. 186 (7), 2164-2172 (2004)

```

```

2 (bases 1 to 301258)
Nascimento,A.L., Ko,A.I., Martins,E.A., Monteiro-Vitorello,C.B.,
Ho,P.L., Haake,D.A., Verjovski-Almeida,S., Hartskeerl,R.A.,
Marques,M.V., Oliveira,M.C., Menck,C.F., Leite,L.C., Carrier,H.,
Coutinho,L.L., Degraeve,W.M., Dellagostin,O.A., El-Dorri,H.,
Ferreto,E.S., Ferro,M.I., Furlan,L.R., Gamberini,M., Gigliotti,E.A.,
Goes-Neto,A., Goldman,G.H., Goldman,M.H., Harakava,R.,
Jeronimo,S.M., Junqueira-De-Azevedo,I.L., Kimura,E.T.,
Kuramae,E.E., Lemos,E.G., Lemos,M.V., Marino,C.L., Nunes,L.R., De
Oliveira,R.C., Pereira,G.G., Reis,M.S., Schriefer,A.,
Siqueira,W.J., Sommer,P., Tsai,S.M., Simpson,A.J., Ferro,J.A.,
Camargo,L.E., Kitajima,J.P., Setubal,J.C. and Van Sluys,M.A.
Comparative Genomics of Two Leptospira interrogans Serovars Reveals
Novel Insights into Physiology and Pathogenesis
J. Bacteriol. 186 (7), 2164-2172 (2004)

```

```

TITLE
JOURNAL
PUBLISHED
REFERENCE
AUTHORS

```

```

Jeronimo,S.M., Junqueira-De-Azevedo,I.L., Kimura,E.T.,
Kuramae,E.E., Lemos,E.G., Lemos,M.V., Marino,C.L., Nunes,L.R., De
Oliveira,R.C., Pereira,G.G., Reis,M.S., Schriefer,A.,
Siqueira,W.J., Sommer,P., Tsai,S.M., Simpson,A.J., Ferro,J.A.,
Camargo,L.E., Kitajima,J.P., Setubal,J.C. and Van Sluys,M.A.
Direct Submission
Submitted (29-FEB-2004) Laboratorio de Bioinformatica/Instituto de
Computacao, Universidade Estadual de Campinas, Avenida Albert
Einstein 1251 Box 6176, Campinas, SP 13084-971, Brasil
Location/Qualifiers
1. 301258
/organism="Leptospira interrogans serovar Copenhageni str.
Fioacruz Li-130"
/mol_type="genomic DNA"
/strain="Fioacruz Li-130"
/serovar="copenhagani"
/db_xref="taxon:267671"
/chromosomes="I"
complement(300..1904)
/locus_tag="LIC12252"
complement(300..1904)
/locus_tag="LIC12252"
/notes="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAS70824.1"
/db_xref="GI:45601343"
/translation="MKNDNQILQRRSIPFCKIRKWNKNKPYKSFFTFVIRPGCIF
ISSLKISAQENFTGCDPEAKRDLPSIDRVKQMKDLDNKKEGYFTGLPLINDP
NEGIGYVRVYGNNGKSDPFETPYVLRFFAQYFNTTKAQHQLSLDMPYVAT
QWRLRADALLTPTTLTVFGVGESTLKLPTHERNQPGAPQVTNGQYQJESFTQYR
PGSGDPIELGRLYSGVSGGFVNTDMYNYTITQPLNFSRSPSPFHGTVRLVA
GRLSNIVHNDGKFKVSDPIFESTPLNSGOVNPNAKLTEDAEAGKILYHGCF
VNTAKGLVIDTRDFEPDPSGVFLEGITYEKASKTIASDFQKYGHAKEFFYSPPK
TFEKLVLASFGFISDGNVPPEFYRNLMTENTVSGLGGLRLRGTKQDRFVGRAMG
FQTLERKWFVSFGVEYFALNVPFWDGFRVWDEHKAGLTDYKYSGGLGURIAWN
QATIMDYAVSRDQKLFVNFSAF"
complement(2436..3119)
/locus_tag="LIC12253"
complement(2436..3119)
/locus_tag="LIC12253"
/notes="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/codon_start=1
/transl_table=11
/product="putative lipoprotein"
/protein_id="AAS70825.1"
/db_xref="GI:45601344"
/translation="MKYQYILIFVILFVFFVNCERKPDPTTIGFNEDEKVLLAGML
FNPYSVDPTSGTITDTSRLMKICTQGAALRVGONGYDCEGINDASTIIGRYGAS
LQYCSLNDNCNTISIPVLGQTFPGFSGTSEAYDTCSDRTAGHSDWRLPTLPELK
ALTAGSLNSFILKPFNTVEDHYDTSWTKEGTVDTHTTISFEASHGFEEMAFKTRN
FVRCVRSRP"
complement(3127..4695)
/locus_tag="LIC12254"
complement(3127..4695)
/locus_tag="LIC12254"
/notes="transmembrane protein; identified by sequence
similarity; putative; ORF located using Blastx/Glimmer"
/codon_start=1
/transl_table=11
/product="outer membrane protein"
/protein_id="AAS70826.1"
/db_xref="GI:45601345"
/translation="MVLKYNRRKMKKFLIFSVFISLFPVTTGAVYGOQDCSKLAF
MDDQVRKPTDLPPFISEMRRLPEDICKKGWFPPTGLPLNSDNPVGVGVVFL
INNGKSDPFEEYAPYRFAQYFNTTKNQYQDISFDAPYIFDTRKLRGLDIYDT
NPNLYYGGESLETLSQERNQSGEVVRNATYHERKNIYFTFPGGPGDPVDFOG
NFTYSGPNDADFRTDRTNIDIRSPQANLSGEHIFFGGLVRAVAGIRVSQNLKIF
DGOQVKSQDPLTEGTTFPSNGITPNKTKVTEADAEAGKIGANGNKVSRVGLVLD
RDLPEDPNRGMFEATYEKAKAFAGSDGFQYSKYFTQVKLFYSPFPKVDKLVLAGRA

```

```

gene
FCLTRGDAPFFPYRNLWSTEGITGLGLRTLRYGKQDRFAGKAMGWNIELRWKFFD
FNIAGOHFALNLVPFMDPGRVWDDHNVGLKDYKYSRGLGFRFIANNQSTILMLDYAVS
KEDKQVFNHNFH"
4726..5001
/locus_tag="LIC12255"
CDS
4726..5001
/locus_tag="LIC12255"
/notes="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAS70827.1"
/db_xref="GI:45601346"
/translacion="MIFQRISDIEIILLFTLGMGFIITRIKIORKAIFITNLGART
LSLNGYFNYSIMIHSSSESDPNFIDRFLKCGNYHKLRFVELILKL"
complement (5668..6327)
/locus_tag="LIC12256"
complement (5668..6327)
/locus_tag="LIC12256"
/notes="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/codon_start=1
/transl_table=11
/product="membrane metalloendopeptidase"
/protein_id="AAS70828.1"
/db_xref="GI:45601347"
/translacion="WRSDFTVIKTKIRKILGHLIVQSLVFTFVLCSENKLEIPKIFE
SRRELGPDSIDKFTSEHNEYVSDGDFDPVGNPAKGYMDQPGFNGHGLGDMNAVQ
RNDYQDPVYSHGLVKGFDGSGPGTVMITLHLLNGKRINSLYAHLSKLNVSXGD
QIRKGMIGRIGDANGRYGPHLHFEMRDDFFLPTGPGYSDPSGLNPKVIRKXKRL
K"
6971..7483
/locus_tag="LIC12257"
gene
6971..7483
/locus_tag="LIC12257"
/notes="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAS70829.1"
/db_xref="GI:45601348"
/translacion="MPKSKTNKKFOIANLLPEHREASIELVNQFRLINSIQLDGI
KIRPRAGTMDIVLKGKTKGVLGLGFLGDELVSLLIARTDEKPYLEEQKTLFIDL
AVTKRKQKSGYMKPLVLACESWAKQDEPKSIELVAIVENAVSFWKAMGYDPFYIR
FRKSPKWDSI"
complement (8468..9919)
/locus_tag="LIC12258"
complement (8468..9919)
/locus_tag="LIC12258"
/notes="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAS70830.1"
/db_xref="GI:45601349"
/translacion="MNRIFRNWNWIGLWILWSLPPIFGEGRLNLPPEARIPLDGKKL
EPSELAKEGWVVTALPLLSSDPVRGQGGIRASLPFNGKKTDLVYEAYRSKLTLL
QLFQTQGVKNHFQFDSFILNTAFRWKSSLSLDYNPNSQYFGIGESSLSLSYSYPR
NLSGIRGVNANFDAYTSQSYIRPRTQSEITPTVSDQYQNLQFNSTFFNGIDYT
FKAWKQVWANEISRNIIIGHSDGWHPSKDPYFAGSIWETPVNGESKLTIDYRAGKI
RGYNGGDIYVFRAGIYDTRDPEPDGRIELAEINVANYSKRTGSDFNKNIFFQTKY
FYKILPDVEELVATRVAGLYTSSCAPFSEVRYMWSLDGPMGTGIGLOTMRGYRODR
FVAPVWFGSAEPRWRAFPKIFDELFTLSLVPFVDVGRWDSKELINLQGYKHSWS
GPRIIWQATVILIDFAKSKDSQMFVDFSHAF"
10798..14472
/locus_tag="LIC12259"
gene
10798..14472
/locus_tag="LIC12259"
/notes="contains alpha integrin like repeat domains;
identified by sequence similarity; putative; ORF located
```

```

using Blastx/Glimmer"
/codon_start=1
/transl_table=11
/product="cytoplasmic membrane protein"
/protein_id="AAS70831.1"
/db_xref="GI:45601350"
/translacion="MISVIKTFISFSLIFGIGCVQCKGMSMPFFAYLDFSNRSADS
SPNYSQISPGPNVTGTSINTSIQVFSQKLDSSSIQSIQLTQCNITIIICGNFTSTEK
TLLNPNSSLAASYVSISKDIKSDMGSSLSRDTWSTNTTIDVLAADVSLRTP
TIGANLVNNTSVQIAFETETWNCISINIVNFKNNVTNLEPNSVCLGSAVLTIPN
NPLAFNTYRVYDILSTAKDLANNPLNAYNMTFTTGVAPELDTVTVSVFVNPVPAQNV
PINETISIAFSEPNICATIIIGSIVLDDNIIIPSGVNGPCTGTATSPGLNLTPTN
NYTIVTSNAITDLQNNPLTPSTWSTFTAAADVQTPVTFTVPSANANGVGTNANPMV
VFSBPMCSASVTSAPFLKROATGYLLIGSVNCFGTSATMTDPDPNPLAFNTTYVEI
DOGALDVFNNPLIPINNFTTGPDPDLTPPSAVVTPANAAGVPTNGGVSIAPSEAM
NCGSLGITLDDDDPTTPTGTVIPINICNGNTVSAFPIPLAENTTVTVTILANTVD
SNNAALNGNYAWSFTTGVAPDLPVQVSLVPSLGAVGVANANITVAFNETINCST
LNTVNGINGTAVNCSSATFIPSAITPLNAGNTYATILITVNDIVNPIGAAFGWS
FTTGAAPDVTPTVVTIQNLRNNSIVESGFVGTATDAGTITSVEVSLDNGAFVATGP
NPMFKLPDINTWKSQHTIIRAKDLANNLTAAISVRKNNKDIINGDYVDLV
SAEYQGLLYIFHSSGNAGMTITNAQSAKIIIVGAEEFGRTVSMGLDNGDGPADVI
SGAPWNGAQRVYIFHSSGNAGVNIISFGFATKTIISANAGARFGDSIVTGLDNGDG
Query Match 72.8%; Score 18.2; DB 1; Length 301258;
Best Local Similarity 87.0%; Pred. No. 8.4e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 AAAAAACACACAGCAGCGCCCTC 23
|||||
Db 117718 AAAAAACGACAAAGCGATCGCTC 117740
RESULT 48
CR378680 349751 bp DNA linear BCT 17-APR-2005
LOCUS Photobacterium profundum SS9 chromosome 2; segment 6/7.
DEFINITION CR378680 CR354532
ACCESSION CR378680.1 GI:46916688
VERSION complete genome.
KEYWORDS Photobacterium profundum SS9
SOURCE Photobacterium profundum SS9
ORGANISM Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
Vibrionaceae; Photobacterium.
REFERENCE
1 Vezzi, A., Campanaro, S., D'Angelo, M., Simonato, F., Vitulo, N.,
Laurio, F., Cesaro, A., Malacrida, G., Simonati, B., Cannata, N.,
Bartlett, D. and Valie, G.
Genome Analysis of Photobacterium profundum reveals the complexity
of high pressure adaptations
Unpublished
2 Cestaro, A.
Direct Submission
Submitted (22-MAR-2004) Cestaro A., CRIBI, Biotechnology Centre,
University of Padova via U. Bassi 58/B, 35131 Padova, Italy
FEATURES
source
1..349751
/organism="Photobacterium profundum SS9"
/mol_type="genomic DNA"
/strain="ss9"
/db_xref="taxon:298386"
/chromosome="2"
/complement (314..1732)
/gene="STM1598"
/locus_tag="PBRB1592"
complement (314..1732)
/locus_tag="PBRB1592"
/gene="STM1598"
/locus_tag="PBRB1592"
/notes="predicted by orpheus program
predicted by glimmer program"
/codon_start=1
/evidence=not_experimental
/transl_table=11
```



```
/product="Hypothetical transcriptional regulator"
/protein_id="CAG23452.1"
/db_xref="GI:46916689"
/db_xref="GOA:Q6LGX8"
/db_xref="InterPro:IPR000524"
/db_xref="UniProt/TREMBL:Q6LGX8"
/translation="MAPQTRRYHAVAITEIQOQIEQRILWLPFGDRIPSRVMTKKHSISP
MTVLKAYELLETEGWYAKPKSGYFVTAHLNRLAEPQOMPOVNNRSIKINAHIFDVL
TACKRPDVPFGSAFPDPLFPLKOLGRALAOCTMTKPAQSTVTDLPNGNIALRRAIA
KRYIRGIDIVITSGAMESLGLSLMAVTPQGDTVIAESPAPFYGKALQAVLRRL
KAVEIATHGMSVNALEAAIKTHDIKACMLSKYONPLGASMPDENKAIANNLLMK
NIPTLEDDYNELYSFSESQPTPIKAYDKRMWLVHCGSFKCLAPGFRVGVVAGQYA
KQIEQLMGTSLTAVPNQLAQYVHLGGYDILHRLRLRQLKARQQOQMLDAIEQYFP
RSTQITRPEGYFLWVLPKHINTSLLQQLLDVHNISITAPGSLFASDDRYKNCMRIN
CSYPWERTQALKTVANCIKK"
complement(2013..3242)
/gene="VP0623"
/locus_tag="PBPRB1593"
complement(2013..3242)
/gene="VP0623"
/locus_tag="PBPRB1593"
/notes="predicted by orpheus program
predicted by glimmer program"
/codon_start=1
/evidence=not experimental
/transl_table=11
/product="hypothetical glycine/D-amino acid oxidase"
/protein_id="CAG23453.1"
/db_xref="GI:46916690"
/db_xref="GOA:Q6LGX7"
/db_xref="InterPro:IPR006076"
/db_xref="UniProt/TREMBL:Q6LGX7"
/translation="MIGLTSAWYLAEGHVTVIDRQSKAETSANAGQISVGYSS
PWAAPGIVKALWLAQHPKIKPSLSDVTVWASMLANQERYQNNKARMLRV
ANYSRCDILDRHQPTNKGQGTQLQVFNAAQIEAVAKDIOVLAESTRYEELDV
AGTAAEPKGVNKLVLGGRLPDDDEGTGCFQOQLTEIAIKAGVKHFNVDVQKL
NOTQGVSSYSTMGALKADAYVMAGSYSSLSPLGLSIPVYPVKYSLTIPVDE
TQAPSVIMDETQYKAVMTREFDQIRVAGTAEAGYDLDLSEQRKATIAMVQDLFNG
GDKSKARYTGLRPMTPDGAPIGKTPFNSLFTNTGHTGLGWTMACGSARLLADVISG
VTPDIPDGUSIARIYA"
3409..4353
/gene="DSDC"
/locus_tag="PBPRB1594"
3409..4353
/gene="DSDC"
/locus_tag="PBPRB1594"
/notes="predicted by orpheus program
predicted by glimmer program"
/codon_start=1
/evidence=not experimental
/transl_table=11
/product="hypothetical transcription regulator LyxR
family"
/protein_id="CAG23454.1"
/db_xref="GI:46916691"
/db_xref="GOA:Q6LGX6"
/db_xref="InterPro:IPR000847"
/db_xref="InterPro:IPR005119"
/db_xref="UniProt/TREMBL:Q6LGX6"
/translation="MLAGLATFFTAARFGSFTQAAEHLHITGTAISOQISHLQBELSF
VLPERHSGIHLTDEGNQLYRVVNSLADIRDVINOIQKAHADGEIHLKLTSPFAFK
WLVPRLHDTQYDPDKIHTFAEGALVDHNDTFDLDAIDYQOTPYICTRNKSQISLIM
AEKLLPVNSPDYKGFDDWNPQDTSQNIKEKVRLLHDAMPWRHAAKDDWCFTFKENG
LTPLNLTSRQHGYNFRNTDMAAAAGLVAARLALVEEDFSKGNLVSPPEPLDA
KAGYLLIQHHSPAIETCFKAWLRQQVNRVYCGSGSGTCL"
5180..5521
/locus_tag="PBPRB1595"
5180..5521
/locus_tag="PBPRB1595"
/notes="predicted by orpheus program
predicted by glimmer program"
/codon_start=1
/evidence=not experimental

/transl_table=11
/product="Hypothetical protein"
/protein_id="CAG23455.1"
/db_xref="GI:46916692"
/db_xref="UniProt/TREMBL:Q6LGX5"
/translation="MLKIYALITIDDLVHSHGISGLSKSYVDHHAHLKVKHVFSETDK
GIVFALVDENSKLHFKLSRPLFNVSRILEVENHRATCFYIIKSMKMTIETEFAD
DRKCLTYSRP"
5810..6286
/locus_tag="PBPRB1596"
5810..6286
/locus_tag="PBPRB1596"
/notes="predicted by orpheus program
predicted by glimmer program"
/codon_start=1
/evidence=not experimental
/transl_table=11
/product="hypothetical protein"
/protein_id="CAG23456.1"
/db_xref="GI:46916693"
/db_xref="UniProt/TREMBL:Q6LGX4"
/translation="MIYKLLPGCIISALLTSTSTLATEISONKHGHPLEYSININ
DSESTAYVTPDMQSDTYGDPILIGSDIQESTVADNPISLI'VQTGHELVVTTITNTKI
NIQARVMKISPLDKTSVQTHIKTLAPDNTSPHTSCQOVFNGEIEIQSVCLASPS"
6876..7454
/locus_tag="PBPRB1597"
6876..7454
/locus_tag="PBPRB1597"
/notes="predicted by orpheus program
predicted by glimmer program"
/codon_start=1
/evidence=not experimental
/transl_table=11
/product="Hypothetical protein"
/protein_id="CAG23457.1"
/db_xref="GI:46916694"
/db_xref="UniProt/TREMBL:Q6LGX3"
/translation="MKPIASDNAQDDGLKHISMNI SAYHSLILREKTKVLKDVL
SALHSAVARTLPOITALTATTTTAAQORVHI VHSMEPEMSECGVSALWSSLPAL
YCLFLAESEGHLSLTDQDIVALLRVDFDKRVLDDTAIEILYRARDHECALAADIL
QQPLLHKEGLIPAIRDCKPIGFESFEDLIARQ"
7512..7847
/locus_tag="PBPRB1598"
7512..7847
/locus_tag="PBPRB1598"
/notes="predicted by orpheus program
predicted by glimmer program"
/codon_start=1
/evidence=not experimental
/transl_table=11
/product="Hypothetical conjugal transfer protein, TrbC"
/protein_id="CAG23458.1"
/db_xref="GI:46916695"
/db_xref="InterPro:IPR007039"
/db_xref="UniProt/TREMBL:Q6LGX2"
/translation="MRCQWRQWSQWLKANIAVFGFWPFLAQASNTGMPYSPLEK
VNSITGPVAFGASVLGIVAAIGLIFGGMNGFIQRLMILSLIVISIVNATNLMSPFL
FNVSSTLV"
7858..8151
/locus_tag="PBPRB1599"
7858..8151
/locus_tag="PBPRB1599"
/notes="predicted by orpheus program
predicted by glimmer program"
/codon_start=1
/evidence=not experimental
/transl_table=11
/product="hypothetical TrbD protein"
/protein_id="CAG23459.1"
```

gene /db_xref="GI:46916696"
/db_xref="InterPro:IPR007909"
/db_xref="UniProt/TREMBL:Q6LGX1"
/translation="WNTDEPLVCRPVFAFNRRSHLFLGGERELVLSMVLICIVLIVALQ
NIYTAMGVVLLIVSITVARMMAKSDPQLTKTYRRFARYQRFYPSNGMKQYGG"
8153..10549

CDS /gene="AGR_PTI"
/locus_tag="PBRB1600"
8153..10549
/gene="AGR_PTI"
/locus_tag="PBRB1600"
/note="predicted by orpheus program
predicted by glimmer program"
/codon_start=1
/evidence=not_experimental

Query Match 72.8%; Score 18.2; DB 1; Length 349751;
Best Local Similarity 87.0%; Pred. No. 8.4e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAACGACACGAGCGCCCTC 23
|||||
Db 4946 AAAAAACTACACGAGCGCCCGC 4968

RESULT 49
AX432478/c
LOCUS AX432478 1368 bp DNA linear PAT 28-JUN-2002
DEFINITION Sequence 893 from Patent WO0229113.
ACCESSION AX432478
VERSION AX432478.1 GI:21657282
KEYWORDS
SOURCE Bacillus licheniformis
ORGANISM Bacillus licheniformis
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
REFERENCE 1
AUTHORS Berka,R. and Clausen,I.G.
TITLES Methods for monitoring multiple gene expression
JOURNAL Patent: WO 0229113-A 893 11-APR-2002;
Novozymes Biotech, Inc. (US); Novozymes A/S (DK)
FEATURES
source
1..1368
/organism="Bacillus licheniformis"
/mol_type="unassigned DNA"
/db_xref="taxon:1402"

ORIGIN

Query Match 71.2%; Score 17.8; DB 6; Length 1368;
Best Local Similarity 90.5%; Pred. No. 1.7e+03;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AAAAAACGACACGAGCGCCCT 22
|||||
Db 1250 AAAAAACGACAAAGGAGCCCT 1230

RESULT 50
AX954557/c
LOCUS AX954557 2347 bp DNA linear PAT 08-JAN-2004
DEFINITION Sequence 23 from Patent WO03093453.
ACCESSION AX954557
VERSION AX954557.1 GI:40783930
KEYWORDS
SOURCE Bacillus licheniformis
ORGANISM Bacillus licheniformis
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
REFERENCE 1
AUTHORS Andersen,J.T., j Rgensen,S.T., Rasmussen,M.D., Olsen,P.B. and
Clausen,I.G.
TITLES Improved bacillus host cell
JOURNAL Patent: WO 03093453-A 23 13-NOV-2003;
Novozymes A/S (DK)
FEATURES
Location/Qualifiers

source

1..2347
/organism="Bacillus licheniformis"
/mol_type="unassigned DNA"
/db_xref="taxon:1402"
513..1859
/note="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAF05474.1"
/db_xref="GI:40783931"
/translation="MISAAASLLLTCTFFPLPSTAQTAIGNETQRLTYPVLTAKTP
EEAGFSKKLKAVDRLIEQDVKAGFPGAALILIKDGKLIKKEVGYKQKNGLIALKH
PKMKANTMFDLASNTMYAVNFALQHLVSTGKLDLNKNIQYLPDFKDHPEDDVKGK
NRLVIDLHNNAGFPASWYDPSAGHLYSQSRKTLVLYVKTPLAYEPTKQIYS
DIDYMLGLIIIEKITNEBLDTFVENRFYRPLGLRHTLFPNLQKGFKRSHFAATERLGN
TRDGTISFPNIRTYTLQGEVHDEKAFYSMEGVSGHAGLFSTVDDAAVLQVLMNGGY
GROHLFSSVISQFTEPSKTNPTTYGLGWRNLNGTDMWFMFGKSHASSKAYGHTGTGTV
TIIDPVYQIGIVLLTNKKHSPVINPKENPNQFEGDEFATGKYGVITAYVEALHHQ"
ORIGIN
Query Match 71.2%; Score 17.8; DB 6; Length 2347;
Best Local Similarity 90.5%; Pred. No. 1.7e+03;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 2 AAAAAACGACACGAGCGCCCT 22
|||||
Db 327 AAAAAACGACAAAGGAGCCCT 307

Search completed: February 3, 2006, 21:26:41
Job time : 877 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 3, 2006, 10:53:33 ; Search time 203.111 Seconds
(without alignments)
820.326 Million cell updates/sec

Title: US-10-719-900-15

Perfect score: 25

Sequence: 1 aaaaacgacacagcgagcctcgg 25

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database :

N Geneseq 21.*

1: geneseqn1980s.*

2: geneseqn1990s.*

3: geneseqn2000s.*

4: geneseqn2001as.*

5: geneseqn2001bs.*

6: geneseqn2002as.*

7: geneseqn2002bs.*

8: geneseqn2003as.*

9: geneseqn2003bs.*

10: geneseqn2003cs.*

11: geneseqn2003ds.*

12: geneseqn2004as.*

13: geneseqn2004bs.*

14: geneseqn2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		DB	ID	Description	
		Match	Length				
1	19.8	79.2	770	6	ABQ26883	Abq26883 Oligonucle	
C	2	19.8	79.2	770	6	ABQ26882	Abq26882 Oligonucle
	3	18.8	75.2	763	6	ABQ76772	Abq76772 Frog embr
C	4	18.6	74.4	378	8	ABQ96483	Abq96483 Rice endo
C	5	17.8	71.2	1368	6	ABQ73602	Abq73602 Bacillus
C	6	17.8	71.2	2347	12	ADQ32053	Adq32053 DNA encou
C	7	17.8	71.2	115780	13	ABQ32610	Abq32610 Mouse cau
8	17.6	70.4	559	6	ABK79534	Abk79534 Bacillus	
9	17.6	70.4	597	6	ABK79471	Abk79471 Bacillus	
10	17.6	70.4	753	6	ABQ39709	Abq39709 Oligonucle	
C	11	17.6	70.4	753	6	ABQ39708	Abq39708 Oligonucle
C	12	17.6	70.4	1017	6	ABQ41506	Abq41506 Oligonucle
13	17.6	70.4	1017	6	ABQ41507	Abq41507 Oligonucle	
14	17.6	70.4	6073	5	ABQ28495	Abq28495 Human pro	
15	17.6	70.4	6073	5	ABV22669	Abv22669 Human pro	
16	17.6	70.4	295772	12	ADQ97433	Adq97433 Human cau	
17	17.2	68.8	1797	13	ADP745037	Adt45037 Bacteria	
18	17.2	68.8	2000	11	ACL38755	ACL38755 Rice str	
C	19	17.2	68.8	6240	10	ADB84162	Adb84162 Gene exp

C	20	17	68.0	283	4	AAF44516
C	21	17	68.0	603	11	ABD00388
C	22	17	68.0	746	3	AAA43897
C	23	17	68.0	3271	4	AAK84896
C	24	17	68.0	4473	13	ADR07098
C	25	17	68.0	6092	5	AAK65021
C	26	17	68.0	7318	14	AAK35989
C	27	17	68.0	16605	4	AAK87049
C	28	17	68.0	16605	4	AAK65489
C	29	17	68.0	17431	4	AAK90339
C	30	17	68.0	17431	4	AAI57710
C	31	17	68.0	17431	6	ABQ99887
C	32	17	68.0	17431	10	ABQ93040
C	33	17	68.0	62230	12	ADQ97885
C	34	17	68.0	62278	13	ABD33618
C	35	17	68.0	86765	10	ADD14752
C	36	17	68.0	115284	11	ACN44296
C	37	17	68.0	163701	13	ABD33351
C	38	17	68.0	163701	14	ADZ13376
C	39	17	68.0	198161	6	ABK83564
C	40	17	68.0	198161	12	ADQ17348
C	41	17	68.0	198161	13	ADR52701
C	42	17	68.0	208765	12	ADQ97430
C	43	17	68.0	215126	12	ADQ97362
C	44	16.8	67.2	2217	12	ADQ23222
C	45	16.6	66.4	354	5	ABV14306
C	46	16.6	66.4	399	5	ABV05137
C	47	16.6	66.4	401	12	ADQ20479
C	48	16.6	66.4	412	13	ACF85290
C	49	16.6	66.4	459	5	ABV44224
C	50	16.6	66.4	459	5	ABQ35393
C	51	16.6	66.4	783	6	ABQ45189
C	52	16.6	66.4	783	6	ABQ45188
C	53	16.6	66.4	936	6	ABQ48607
C	54	16.6	66.4	936	6	ABQ48606
C	55	16.6	66.4	1243	13	ADX11936
C	56	16.6	66.4	1252	13	ADX64841
C	57	16.6	66.4	1488	4	ABL16770
C	58	16.6	66.4	1705	13	ADX60957
C	59	16.6	66.4	2000	11	ACL35569
C	60	16.6	66.4	2022	8	ABZ36091
C	61	16.6	66.4	2025	2	AAK81876
C	62	16.6	66.4	2025	3	AAK88684
C	63	16.6	66.4	2025	4	AAI11128
C	64	16.6	66.4	2025	14	ADZ14867
C	65	16.6	66.4	15118	10	ADG87690
C	66	16.6	66.4	31312	13	ABD33526
C	67	16.6	66.4	81684	13	ABD33502
C	68	16.6	66.4	106315	11	ACN43966
C	69	16.6	66.4	110000	11	ACN44150_2
C	70	16.4	65.6	1343	6	ABZ78056
C	71	16.4	65.6	3259	6	ABZ78057
C	72	16.2	64.8	345	6	ABQ90353
C	73	16.2	64.8	428	4	AAK74513
C	74	16.2	64.8	452	3	AAZ36720
C	75	16.2	64.8	571	6	ABQ47739
C	76	16.2	64.8	571	6	ABQ47738
C	77	16.2	64.8	579	13	ADQ53456
C	78	16.2	64.8	607	6	ABQ48185
C	79	16.2	64.8	607	6	ABQ48184
C	80	16.2	64.8	707	6	ABQ36974
C	81	16.2	64.8	707	6	ABQ36975
C	82	16.2	64.8	776	6	ABQ28458
C	83	16.2	64.8	776	6	ABQ28459
C	84	16.2	64.8	882	2	AAQ79605
C	85	16.2	64.8	904	6	ABQ28483
C	86	16.2	64.8	904	6	ABQ28482
C	87	16.2	64.8	1002	6	ABQ90354
C	88	16.2	64.8	1012	6	ABQ35513
C	89	16.2	64.8	1012	6	ABQ35512
C	90	16.2	64.8	1017	6	ABQ39579
C	91	16.2	64.8	1017	6	ABQ39578
C	92	16.2	64.8	1061	6	ABQ21136

AAf44516	Mouse dex
Abd00388	Klebsiell
Aaa43897	Human sec
Aak84896	Human imm
Adi07098	Full leng
Aas65021	DNA encod
Aas5989	Long term
Aak87049	Human imm
Aak65489	Human dig
Aak90339	Human col
Aai57710	Human col
Abq99887	Genomic D
Adq93040	Human col
Adq97885	Mouse can
Abd33618	Murine ca
Adi14752	Human sec
Acn44296	Mouse gen
Abd33351	Murine ca
Adz13376	Murine ca
Abk83564	Human cdn
Adi17348	Human sof
Adi52701	Drug ther
Adq97430	Mouse can
Adq97362	Mouse can
Adq23222	Human sof
Abv14306	Human pro
Abv05137	Human pro
Adq20479	Human sof
Acf85290	Human SIR
Abv44224	Human pro
Abv35393	Human pro
Abq45189	Oligonucl
Abq45188	Oligonucl
Abq48607	Oligonucl
Abq48606	Oligonucl
Adx11936	Plant ful
Adx64841	Plant ful
Abi16770	Drosophil
Adi60957	Plant ful
Adi35569	Rice stre
Adi35569	Rice stre
Aax81876	DNA encod
Aaa88684	Sweetgum
Adi1128	L. styrac
Adi14867	Syringyl
Adc87690	Human mam
Abd33526	Murine ca
Abd33502	Murine ca
Acn43966	Human gen
Continuation	(3 of
Abz78056	Human bre
Abz78057	Human bre
Abq90353	M. capsul
Aak74513	Human imm
Aaz36720	Nucleotid
Abq47739	Oligonucl
Abq47738	Oligonucl
Adq53456	Novel can
Abq48185	Oligonucl
Abq48184	Oligonucl
Abq36974	Oligonucl
Abq36975	Oligonucl
Abq28458	Oligonucl
Abq28459	Oligonucl
Abq79605	Yeast cyt
Abq28483	Oligonucl
Abq28482	Oligonucl
Abq90354	M. capsul
Abq35513	Oligonucl
Abq35512	Oligonucl
Abq39579	Oligonucl
Abq39578	Oligonucl
Abq21136	Oligonucl

93	16.2	64.8	1061	6	ABQ21137	Abq21137 Oligonucleotide
94	16.2	64.8	1409	1	AA50474	Sequence
95	16.2	64.8	2013	3	AA229206	Aaz229206 Rat myel
96	16.2	64.8	2013	3	ABN86599	Abn86599 Rat myel
97	16.2	64.8	2185	10	ADD48351	Add48351 Rat gene
98	16.2	64.8	2426	6	ABQ70640	Abq70640 Listeria
99	16.2	64.8	2468	12	ADP72650	Adp72650 Renal tox
100	16.2	64.8	2468	13	ADP72650	Adp72650 Rat cDNA
101	16.2	64.8	2474	10	ABT41975	Abt41975 Toxicity
102	16.2	64.8	2475	9	ACF25356	Acf25356 Rat MAG g
103	16.2	64.8	2475	10	ADB58278	Adb58278 Toxicity
104	16.2	64.8	2475	12	ADR13958	Adr13958 Rat myel
105	16.2	64.8	2475	13	ADT89534	Adt89534 Rattus no
106	16.2	64.8	2475	13	ADG99346	Adg99346 Rat MAG p
107	16.2	64.8	2826	14	ACL67183	Acl67183 M. xanthu
108	16.2	64.8	27786	4	ABL27786	Abi27786 drosophil
109	16.2	64.8	5975	6	ABN80065	Abn80065 Human che
110	16.2	64.8	5618	4	ABL27786	Abi27786 Human car
111	16.2	64.8	15425	4	AA336154	Aas336154 Human car
112	16.2	64.8	15425	10	ADG46848	Adg46848 Human car
113	16.2	64.8	15425	13	ADJ08266	Adj08266 Human car
114	16.2	64.8	16563	14	ACL65537	Acl65537 M. xanthu
115	16.2	64.8	16844	14	ACL64713	Acl64713 M. xanthu
116	16.2	64.8	17789	4	ABL07956	Abi07956 Drosophil
117	16.2	64.8	28493	14	ACL64778	Acl64778 M. xanthu
118	16.2	64.8	41322	9	AA162633	Aal62633 Human CD3
119	16.2	64.8	60335	13	ACN37224	Acn37224 Human per
120	16.2	64.8	104096	13	ADG56092	Adg56092 Streptom
121	16.2	64.8	106938	13	ABD33432	Abd33432 Human can
122	16.2	64.8	106938	13	ADR67034	Adr67034 Human can
123	16.2	64.8	107310	14	ADZ13456	Adz13456 Human can
124	16.2	64.8	110000	6	ABQ30341_20	Continuation (21 o
125	16.2	64.8	263852	13	ADG99460	Adg99460 Murine ki
126	16.2	64.8	340449	8	AAU52198	Aal52198 Human sec
127	16.2	64.8	184	3	AA12801	Aac12801 Human sec
128	16.2	64.8	248	5	AA897656	Aas897656 DNA encod
129	16.2	64.8	283	6	ABL66260	Abi66260 Lung canc
130	16.2	64.8	364	5	AA80289	Aas80289 DNA encod
131	16.2	64.8	420	3	AAH30228	Aah30228 Human col
132	16.2	64.8	432	8	ABX62004	Abx62004 Arabidops
133	16.2	64.8	432	10	ABX61445	Abx61445 Arabidops
134	16.2	64.8	456	5	AA911175	Aas911175 DNA encod
135	16.2	64.8	459	5	AA91010	Aas91010 DNA encod
136	16.2	64.8	502	6	ABQ50216	Abq50216 Oligonucle
137	16.2	64.8	502	6	ABQ50217	Abq50217 Oligonucle
138	16.2	64.8	531	3	AAF09718	Aaf09718 Fusarium
139	16.2	64.8	531	13	ADU53759	Adu53759 Fusarium
140	16.2	64.8	531	14	ADZ91762	Adz91762 Fusarium
141	16.2	64.8	542	6	ABQ34321	Abq34321 Oligonucle
142	16.2	64.8	542	6	ABQ34320	Abq34320 Oligonucle
143	16.2	64.8	559	6	ABQ38716	Abq38716 Oligonucle
144	16.2	64.8	559	6	ABQ38717	Abq38717 Oligonucle
145	16.2	64.8	587	6	ABQ40361	Abq40361 Oligonucle
146	16.2	64.8	587	6	ABQ40360	Abq40360 Oligonucle
147	16.2	64.8	593	6	ABN23343	Abn23343 Human ORF
148	16.2	64.8	597	6	ABQ15808	Abq15808 Oligonucle
149	16.2	64.8	597	6	ABQ15809	Abq15809 Oligonucle
150	16.2	64.8	597	6	ABQ40253	Abq40253 Oligonucle
			597	6	ABQ40252	Abq40252 Oligonucle

ALIGNMENTS

RESULT 1
ABQ26883
ID ABQ26883 standard; DNA; 770 BP.

XX ABQ26883;

XX 12-JUL-2002 (first entry)

XX Oligonucleotide for detecting cytosine methylation SEQ ID NO 13474.

KW Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis; drug; side effect; cancer; central nervous system; cardiovascular; gastrointestinal; respiratory system; single nucleotide polymorphism; SNP; cell differentiation; ds.

OS Homo sapiens.

XX WO200218632-A2.

PN WO200218632-A2.

XX 07-MAR-2002.

XX 01-SEP-2001; 2001WO-EP010074.

XX 01-SEP-2000; 2000DE-01043826.

XX 05-SEP-2000; 2000DE-01044543.

XX (EPIG-) EPIGENOMICS AG.

XX Olek A, Piepenbrock C, Berlin K, Guetig D;

XX WPI; 2002-371829/40.

XX Determining the degree of cytosine methylation in genomic DNA, useful for diagnosis and prognosis, comprises selective hybridization of amplicons from chemically treated DNA.

XX Claim 12; 56pp + Sequence Listing; 56pp; German.

XX This invention describes a novel method for determining the degree of methylation of a particular cytosine in a motif 5'-CpG-3', present in a genomic sample of DNA. The sample is treated chemically to convert cytosine (C) but not methylated C, to uracil, then part of the genomic DNA that contains the target C is amplified to form a labeled amplicon. The amplicon is hybridised to two classes, each with at least one member, of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the degree of hybridisation to both classes is determined from the label on the amplicon. From the ratio of labels hybridised to the two classes of oligomers, the degree of methylation is calculated. The method is used: (i) for diagnosis and/or prognosis of side effects of therapeutic drugs and of a wide range of diseases, e.g. cancer, disorders of the central nervous, cardiovascular, gastrointestinal and respiratory systems etc., particularly by detecting mutations or single nucleotide polymorphisms (SNP's); and (ii) for differentiation of cell or tissue types and for investigating cell differentiation. The method allows the methylation status of many C residues to be determined simultaneously. ABQ1410-ABQ54121 represent genomic DNA sequences used to illustrate the method for determining the degree of cytosine methylation described in the disclosure of the invention

XX Sequence 770 BP; 258 A; 318 C; 85 G; 109 T; 0 U; 0 Other;

XX Query Match 79.2%; Score 19.8; DB 6; Length 770;

XX Best Local Similarity 91.3%; Pred. No. 29;

XX Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAACGACACGAGCCCTC 23

DB 362 AAAAAACGACACGAGCCCTC 384

RESULT 2

ABQ26882/c

ID ABQ26882 standard; DNA; 770 BP.

XX ABQ26882;

XX 12-JUL-2002 (first entry)

XX Oligonucleotide for detecting cytosine methylation SEQ ID NO 13473.

KW Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis; drug; side effect; cancer; central nervous system; cardiovascular; gastrointestinal; respiratory system; single nucleotide polymorphism; SNP; cell differentiation; ds.

KW SNP; cell differentiation; ds.

OS Homo sapiens.

XX W0200218632-A2.

XX PD 07-MAR-2002.

XX PF 01-SEP-2001; 2001WO-EP010074.

XX PR 01-SEP-2000; 2000DE-01043826.

XX PR 05-SEP-2000; 2000DE-01044543.

XX PA (EPIC-) EPIGENOMICS AG.

XX PI Olek A, Piepenbrock C, Berlin K, Guetig D;

XX DR WPI; 2002-371829/40.

XX PT Determining the degree of cytosine methylation in genomic DNA, useful for
PT diagnosis and prognosis, comprises selective hybridization of amplicons
PT from chemically treated DNA.

XX PS Claim 12; 56pp + Sequence Listing; 56pp; German.

XX CC This invention describes a novel method for determining the degree of
CC methylation of a particular cytosine in a motif 5'-CG-3', present in a
CC genomic sample of DNA. The sample is treated chemically to convert
CC cytosine (C) but not methylated C, to uracil, then part of the genomic
CC DNA that contains the target C is amplified to form a labeled amplicon.
CC The amplicon is hybridised to two classes, each with at least one member,
CC of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the
CC degree of hybridisation to both classes is determined from the label on
CC the amplicon. From the ratio of labels hybridised to the two classes of
CC oligomers, the degree of methylation is calculated. The method is used:
CC (i) for diagnosis and/or prognosis of side effects of therapeutic drugs
CC and of a wide range of diseases, e.g. cancer, disorders of the central
CC nervous, cardiovascular, gastrointestinal and respiratory systems etc.,
CC particularly by detecting mutations or single nucleotide polymorphisms
CC (SNP's); and (ii) for differentiation of cell or tissue types and for
CC investigating cell differentiation. The method allows the methylation
CC status of many C residues to be determined simultaneously. ABQ13410-
CC ABQ54121 represent genomic DNA sequences used to illustrate the method
CC for determining the degree of cytosine methylation described in the
CC disclosure of the invention

XX SQ Sequence 770 BP; 109 A; 85 C; 318 G; 258 T; 0 U; 0 Other;

Query Match 79.2%; Score 19.8; DB 6; Length 770;

Best Local Similarity 91.3%; Pred. No. 29;

Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AAAAAACGACACGCGGCCCTC 23

Db 409 AAAAAACGACACGCGGCCCTC 387

RESULT 3

ABST76772

ID ABST76772 standard; cDNA; 763 BP.

XX AC ABST76772;

XX DT 12-DEC-2002 (first entry)

XX DE Frog embryonic gene sequence Q9924949.

XX KW Frog; ss; embryonic development; developmental disorder; microarray;
cell differentiation.

XX OS Xenopus laevis.

XX PN US2002081610-A1.

XX 27-JUN-2002.

XX 23-JUL-2001; 2001US-00910943.

XX 21-JUL-2000; 2000US-0219658P.

XX (UYRQ) UNIV ROCKEFELLER.

XX Hemmati-Brivanlou A, Altmann CR;

XX WPI; 2002-626534/67.

XX Nucleic acid array containing Xenopus embryonic nucleic acids is useful
PT to identify genes involved in embryonic development, to identify
PT different types of embryonic cells, and to diagnose developmental
PT disorders.

XX Claim 1; Page 42; 823pp; English.

XX CC The invention relates to a nucleic acid array, where each coordinate
CC contains a single nucleic acid species having one of 770 nucleotide
CC sequences (appearing as ABS76747-ABS77516) a of a Xenopus embryonic gene
CC product, or its complement or hybridisable fragment of not less than 20
CC contiguous nucleotides of one of those sequences. Also included are
CC detecting differential expression of embryonic genes, comprising: (a)
CC contacting a nucleic acid array comprising genes expressed in embryonic
CC but not mature cells with nucleic acids from sample and control cells;
CC and (b) detecting differential hybridisation of nucleic acids from the
CC sample cells relative to the control cells; and detecting defects in
CC development, comprising: (a) contacting nucleic acids from test cells
CC undergoing development with a nucleic acid array of gene products known
CC to play a fundamental role in the development process; and (b) detecting
CC a difference in expression of a fundamental gene in the sample cells
CC relative to a standard. The invention is useful to identify genes
CC involved in embryonic development and related processes such as cell
CC differentiation. This would be useful for diagnosing developmental
CC disorders and for identifying different types of embryonic cells. The
CC present sequence is one of the 770 Xenopus embryonic cDNA sequences

XX SQ Sequence 763 BP; 181 A; 223 C; 208 G; 141 T; 0 U; 10 Other;

Query Match 75.2%; Score 18.8; DB 6; Length 763;

Best Local Similarity 90.9%; Pred. No. 84;

Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 AAACGACACGACGCGCCCTCGG 25

Db 549 AAACGACACGACGCGCCCTCGG 570

RESULT 4

ABX96483/c

ID ABX96483 standard; cDNA; 378 BP.

XX AC ABX96483;

XX DT 13-MAY-2003 (first entry)

XX DE Rice endosperm expression sequence label #24.

XX KW Rice; endosperm expression sequence label; gene chip; babel technique;
microarray; ss.

XX OS Oryza sp.

XX PN CN1366050-A.

XX PD 28-AUG-2002.

XX 31-OCT-2001; 2001CN-00135874.

XX 31-OCT-2001; 2001CN-00135874.

XX (UYZH-) UNIV ZHEJIANG.
XX Dong H, Li D;
XX WPI; 2003-240400/24.
XX Rice endosperm expression sequence label and genechip prepared from it.
XX Claim 1; Page 19 (Disclosure); 29pp; Chinese.
XX The invention describes a novel rice endosperm expression sequence label
CC and the gene chip formed from it. The expression sequence label technique
CC is used to create a rice endosperm cDNA library. Non-redundant expression
CC sequence labels are then combined and used to create a gene chip by
CC microarray techniques. This sequence represents a rice endosperm
CC expression sequence
XX Sequence 378 BP; 90 A; 119 C; 87 G; 82 T; 0 U; 0 Other;
SQ Query Match 74.4%; Score 18.6; DB 8; Length 378;
Best Local Similarity 84.0%; Pred. No. 96;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 AAAAAACGACACAGCGCCCTCGG 25
DB 235 AAAAAACGACACAGGGGCTCTGGG 211
RESULT 5
ABK73602/c
ID ABK73602 standard; DNA; 1368 BP.
XX AC ABK73602;
XX 13-AUG-2002 (first entry)
XX Bacillus licheniformis genomic sequence tag (GST) #893.
XX Differential gene expression; genomic sequenced tag; GST;
KW altered culture condition; environmental stress;
KW physiological provocation; ds.
XX Bacillus licheniformis.
XX WO200229113-A2.
XX 11-APR-2002.
XX 05-OCT-2001; 2001WO-US031437.
XX 06-OCT-2000; 2000US-00680598.
XX 27-MAR-2001; 2001US-0279526P.
XX (NOVO) NOVOZYMES BIOTECH INC.
XX (NOVO) NOVOZYMES AS.
XX Berka R, Clausen IG;
XX WPI; 2002-416684/44.
XX Monitoring differential expression of several genes in first Bacillus
PT cell relative to expression of same genes in one or more second Bacillus
PT cells, by using substrate containing Bacillus genomic sequenced tag
PT array.
XX Claim 4; SEQ ID NO 893; 200pp; English.
XX The invention describes a method of monitoring differential expression of
CC genes in a first Bacillus cell relative to expression of the genes in
CC other Bacillus cells, comprising hybridising labelled nucleic acid probes
CC isolated from Bacillus cells to a substrate containing array of Bacillus
CC genomic sequenced tags (GST), examining the array, and determining
XX relative gene expression by an observed hybridisation reporter signal of
a spot in the array. The method is useful for measuring the expression of
genes in a first Bacillus cell relative to expression of the same genes
in one or more second Bacillus cells. The method is useful for monitoring
global expression of several genes from a Bacillus cell, discovering new
genes, identifying possible functions of unknown open reading frames and
monitoring gene copy number variation and stability. Monitoring changes
in expression of genes may be used to provide a representation of the way
in which Bacillus cells adapt to changes in culture conditions,
environmental stress or other physiological provocation. Extensive follow
up characterisation is unnecessary, when one spot on an array equals one
gene or one open reading frame, since sequence information is available.
This sequence represents a genomic sequence tag (GST) used in the method
of the invention. Note: The sequence data for this patent did not form
part of the printed specification, but was obtained in electronic format
directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX Sequence 1368 BP; 314 A; 365 C; 329 G; 360 T; 0 U; 0 Other;
SQ Query Match 71.2%; Score 17.8; DB 6; Length 1368;
Best Local Similarity 90.5%; Pred. No. 2.6e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 2 AAAAAACGACACAGCGCCCT 22
DB 1250 AAAAAACGACAAAGGGAGCCCT 1230
RESULT 6
ADG32053/c
ID ADG32053 standard; DNA; 2347 BP.
XX AC ADG32053;
XX 26-FEB-2004 (first entry)
XX DNA encoding a mutant B_licheniformis secreted polypeptide SeqID 23.
XX mutant; host cell; production yield; shelf life; product stability;
KW purity; secreted; gene; ds.
XX Synthetic.
XX Bacillus licheniformis.
XX WO2003093453-A2.
XX 13-NOV-2003.
XX 25-MAR-2003; 2003WO-DK000198.
XX 10-APR-2002; 2002DK-00000534.
XX (NOVO) NOVOZYMES AS.
XX Andersen JT, Joergensen ST, Rasmussen MD, Olsen PB, Clausen IG;
XX WPI; 2004-053045/05.
XX P-PSDB; ADG32054.
XX New mutant Bacillus licheniformis host cell secreting 5 % less of one or
PT more secreted polypeptides than the parent host cell, useful for
PT producing a product of interest e.g. polypeptides, amino acids or
PT carbohydrates.
XX Disclosure; SEQ ID NO 23; 422pp; English.
XX This invention relates to a novel Bacillus licheniformis (B.
CC licheniformis) mutant host cell derived from a parent B. licheniformis
CC host cell that is mutated in genes encoding secreted polypeptides. host
CC Specifically, it refers to the generation of an improved Bacillus host
CC that reduces the need for product purification caused by contaminant
CC secreted native polypeptides in the culture medium. Accordingly, the
CC present invention describes reducing the expression of these native

CC proteins (e.g. proteolytic enzymes, nutrient uptake factors and signal
CC molecules), which in turn makes it easier to purify the heterologous
CC product of interest and therefore improving the production process.
CC Further benefits of a mutated host cell include an increase in total
CC production yield and a longevity of shelf life attributable to improved
CC product stability and purity. This polynucleotide is a DNA sequence
CC encoding a mutant B. licheniformis secreted polypeptide of the invention.

XX Sequence 2347 BP; 681 A; 567 C; 542 G; 557 T; 0 U; 0 Other;

Query Match 71.2%; Score 17.8; DB 12; Length 2347;
Best Local Similarity 90.5%; Pred. No. 2.8e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AAAAAGCAGCAGCGAGCCCT 22

Db 327 AAAAAGCAGCAGCGAGCCCT 307

RESULT 7

ABD32610/C

ID ABD32610 standard; DNA; 115780 BP.

XX AC ABD32610;

XX DT 18-NOV-2004 (first entry)

XX DE Mouse cancer-associated genomic DNA MD12-017.

XX KW Mouse; ds; cancer-associated protein; gene; cytostatic; cancer;

XX KW Leukaemia; lymphoma; CAP.

XX OS Mus musculus.

XX PN WO2004074320-A2.

XX PD 02-SEP-2004.

XX PF 17-FEB-2004; 2004WO-US004730.

XX PR 14-FEB-2003; 2003US-00367094.

XX PR 14-MAR-2003; 2003US-00388838.

XX PR 15-APR-2003; 2003US-00417375.

XX PR 13-JUN-2003; 2003US-00461862.

XX PR 15-SEP-2003; 2003US-00663431.

XX PR 15-DEC-2003; 2003US-00737318.

XX PA (SAGR-) SAGRES DISCOVERY INC.

XX PI Morris DW, Morris DW, Malandro MS;

XX DR WPI; 2004-652914/63.

XX PT New isolated cancer-associated polynucleotides and polypeptides useful

XX PT for diagnosing, preventing or treating cancers, especially lymphoma and

XX PT leukemia, or in screening for agents that modulate cancer.

XX PS disclosure; seqid 122; 310pp; English.

XX CC The invention relates to an isolated nucleic acid comprising at least 10

XX CC contiguous nucleotides of any of the 233 polynucleotide sequences given

XX CC in the specification, or its complement. The nucleic acids encode cancer-

XX CC associated proteins. Also included are an expression vector comprising

XX CC the isolated nucleic acid cited above, a host cell comprising the above

XX CC recombinant nucleic acid or expression vector, a microarray for detecting

XX CC a cancer-associated (CA) nucleic acid comprising at least one probe

XX CC comprising at least 10 contiguous nucleotides of any of the above-

XX CC mentioned nucleotide sequences, an isolated polypeptide (encoded within

XX CC an open reading frame of a CA sequence selected from any of the 95

XX CC polynucleotide sequences as mentioned in the specification, or its

XX CC complement), an isolated antibody, (or its antigen binding fragment) that

XX CC binds to the above polypeptide, a hybridoma that produces the above

XX CC monoclonal antibody, a pharmaceutical composition comprising the above

CC antibody and a pharmaceutical excipient, a kit for detecting cancer
CC cells(comprising the antibody cited above, methods for diagnosing cancer
CC or for detecting the presence or absence of cancer cells in an
CC individual, a method for inhibiting growth of cancer cells in an
CC individual, a method for delivering a therapeutic agent to cancer cells
CC in an individual, an electronic library comprising the above
CC polynucleotide or polypeptide (or their fragments), methods of screening
CC for anticancer activity or for a bioactive agent capable of modulating
CC the activity of a CA protein (CAP), methods for detecting cancer
CC associated with expression of a polypeptide in a test cell sample, a
CC method for treating cancers and a method for inhibiting the expression of
CC CA gene in a cell. The composition and methods are useful for detecting,
CC diagnosing, preventing and treating cancers, especially lymphoma and
CC leukaemia. These may also be used in screening for agents that modulate
CC cancer. The present sequence is a mouse CAP genomic sequence. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 115780 BP; 27988 A; 26066 C; 28294 G; 33243 T; 0 U; 189 Other;

Query Match 71.2%; Score 17.8; DB 13; Length 115780;

Best Local Similarity 90.5%; Pred. No. 4.5e+02;

Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 AACGACACAGCGAGCCCTCGG 25

Db 111059 AAGGACAAAGCGAGCCCTCGG 111039

RESULT 8

ABK79534

ID ABK79534 standard; DNA; 559 BP.

XX AC ABK79534;

XX DT 13-AUG-2002 (first entry)

XX DE Bacillus clausii genomic sequence tag (GST) #2377.

XX KW Differential gene expression; genomic sequenced tag; GST;

XX KW altered culture condition; environmental stress;

XX KW physiological provocation; ds.

XX OS Bacillus clausii.

XX PN WO200229113-A2.

XX PD 11-APR-2002.

XX PF 05-OCT-2001; 2001WO-US031437.

XX PR 06-OCT-2000; 2000US-00680598.

XX PR 27-MAR-2001; 2001US-0279526P.

XX PA (NOVO) NOVOZYMES BIOTECH INC.

XX PA (NOVO) NOVOZYMES AS.

XX PI Berka R, Clausen IG;

XX DR WPI; 2002-416684/44.

XX PT Monitoring differential expression of several genes in first Bacillus

XX PT cell relative to expression of same genes in one or more second Bacillus

XX PT cells, by using substrate containing Bacillus genomic sequenced tag

XX PT array.

XX PS Claim 11; SEQ ID NO 6825; 200pp; English.

XX CC The invention describes a method of monitoring differential expression of

XX CC genes in a first Bacillus cell relative to expression of the genes in

XX CC other Bacillus cells, comprising hybridising labelled nucleic acid probes

XX CC isolated from Bacillus cells to a substrate containing array of Bacillus

CC genomic sequenced tags (GST), examining the array, and determining
 CC relative gene expression by an observed hybridisation reporter signal of
 CC a spot in the array. The method is useful for measuring the expression of
 CC genes in a first Bacillus cell relative to expression of the same genes
 CC in one or more second Bacillus cells. The method is useful for monitoring
 CC global expression of several genes from a Bacillus cell, discovering new
 CC genes, identifying possible functions of unknown open reading frames and
 CC monitoring gene copy number variation and stability. Monitoring changes
 CC in expression of genes may be used to provide a representation of the way
 CC in which Bacillus cells adapt to changes in culture conditions,
 CC environmental stress or other physiological provocation. Extensive follow
 CC up characterisation is unnecessary, when one spot on an array equals one
 CC gene or one open reading frame, since sequence information is available.
 CC This sequence represents a genomic sequence tag (GST) used in the method
 CC of the invention. Note: The sequence data for this patent did not form
 CC part of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 XX
 SQ Sequence 559 BP; 128 A; 132 C; 142 G; 157 T; 0 U; 0 Other;

Query Match 70.4%; Score 17.6; DB 6; Length 559;
 Best Local Similarity 83.3%; Pred. No. 2.9e+02;
 Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAACGACACAGCGAGCCCTCG 24
 |||||
 Db 203 AAAAAACGACACATCGAAGCTTCG 226

RESULT 9
 ID ABK79471
 XX ABK79471 standard; DNA; 597 BP.
 AC ABK79471;
 XX
 DT 13-AUG-2002 (first entry)
 XX
 DE Bacillus clausii genomic sequence tag (GST) #2314.
 XX
 KW Differential gene expression; genomic sequenced tag; GST;
 KW altered culture condition; environmental stress;
 KW physiological provocation; ds.
 XX
 OS Bacillus clausii.
 XX
 PN WO200229113-A2.
 XX
 PD 11-APR-2002.
 XX
 PF 05-OCT-2001; 2001WO-US031437.
 XX
 PR 06-OCT-2000; 2000US-00680598.
 PR 27-MAR-2001; 2001US-0279526P.
 XX
 PA (NOVO) NOVOZYMES BIOTECH INC.
 PA (NOVO) NOVOZYMES AS.
 XX
 PI Berka R, Clausen IG;
 XX
 XX WPI; 2002-416684/44.

XX Monitoring differential expression of several genes in first Bacillus
 PT cell relative to expression of same genes in one or more second Bacillus
 PT cells, by using substrate containing Bacillus genomic sequenced tag
 PT array.
 XX
 XX Claim 11; SEQ ID NO 6762; 200pp; English.
 PS
 CC The invention describes a method of monitoring differential expression of
 CC genes in a first Bacillus cell relative to expression of the genes in
 CC other Bacillus cells, comprising hybridising labelled nucleic acid probes
 CC isolated from Bacillus cells to a substrate containing array of Bacillus
 CC genomic sequenced tags (GST), examining the array, and determining

CC relative gene expression by an observed hybridisation reporter signal of
 CC a spot in the array. The method is useful for measuring the expression of
 CC genes in a first Bacillus cell relative to expression of the same genes
 CC in one or more second Bacillus cells. The method is useful for monitoring
 CC global expression of several genes from a Bacillus cell, discovering new
 CC genes, identifying possible functions of unknown open reading frames and
 CC monitoring gene copy number variation and stability. Monitoring changes
 CC in expression of genes may be used to provide a representation of the way
 CC in which Bacillus cells adapt to changes in culture conditions,
 CC environmental stress or other physiological provocation. Extensive follow
 CC up characterisation is unnecessary, when one spot on an array equals one
 CC gene or one open reading frame, since sequence information is available.
 CC This sequence represents a genomic sequence tag (GST) used in the method
 CC of the invention. Note: The sequence data for this patent did not form
 CC part of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 XX
 SQ Sequence 597 BP; 141 A; 143 C; 152 G; 161 T; 0 U; 0 Other;

Query Match 70.4%; Score 17.6; DB 6; Length 597;
 Best Local Similarity 83.3%; Pred. No. 3e+02;
 Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAACGACACAGCGAGCCCTCG 24
 |||||
 Db 243 AAAAAACGACACATCGAAGCTTCG 266

RESULT 10
 ID ABQ39709
 XX ABQ39709 standard; DNA; 753 BP.
 AC ABQ39709;
 XX

DT 12-JUL-2002 (first entry)

XX
 DE Oligonucleotide for detecting cytosine methylation SEQ ID NO 26300.
 XX
 KW Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
 KW drug; side effect; cancer; central nervous system; cardiovascular;
 KW gastrointestinal; respiratory system; single nucleotide polymorphism;
 KW SNP; cell differentiation; ds.
 XX

OS Homo sapiens.

XX WO200218632-A2.

XX 07-MAR-2002.

XX 01-SEP-2001; 2001WO-EP010074.

XX 01-SEP-2000; 2000DE-01043826.

XX 05-SEP-2000; 2000DE-01044543.

XX (EPITG-) EPIGENOMICS AG.

XX Olek A, Piepenbrock C, Berlin K, Guetig D;

XX WPI; 2002-371829/40.

XX Determining the degree of cytosine methylation in genomic DNA, useful for
 PT diagnosis and prognosis, comprises selective hybridization of amplicons
 PT from chemically treated DNA.

XX Claim 12; 56pp + Sequence Listing; 56pp; German.

XX This invention describes a novel method for determining the degree of
 CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a
 CC genomic sample of DNA. The sample is treated chemically to convert
 CC cytosine (C) but not methylated C, to uracil, then part of the genomic
 CC DNA that contains the target C is amplified to form a labeled amplicon.
 CC The amplicon is hybridised to two classes, each with at least one member,
 CC of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the

CC degree of hybridisation to both classes is determined from the label on
 CC the amplicon. From the ratio of labels hybridised to the two classes of
 CC oligomers, the degree of methylation is calculated. The method is used:
 CC (i) for diagnosis and/or prognosis of side effects of therapeutic drugs
 CC and of a wide range of diseases, e.g. cancer, disorders of the central
 CC nervous, cardiovascular, gastrointestinal and respiratory systems etc.,
 CC particularly by detecting mutations or single nucleotide polymorphisms
 CC (SNP's); and (ii) for differentiation of cell or tissue types and for
 CC investigating cell differentiation. The method allows the methylation
 CC status of many C residues to be determined simultaneously. ABQ13410-
 CC ABQ54121 represent genomic DNA sequences used to illustrate the method
 CC for determining the degree of cytosine methylation described in the
 CC disclosure of the invention

SQ Sequence 753 BP; 273 A; 303 C; 111 G; 66 T; 0 U; 0 Other;
 Query Match 70.4%; Score 17.6; DB 6; Length 753;
 Best Local Similarity 83.3%; Pred. No. 3e+02;
 Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAACGACACAGCGCCCTCG 24
 |||||
 Db 286 AAAAAACGACCGCGACCCCGCG 309

RESULT 11
 ABQ39708/c
 ID ABQ39708 standard; DNA; 753 BP.

XX AC ABQ39708;

DT 12-JUL-2002 (first entry)

DE Oligonucleotide for detecting cytosine methylation SEQ ID NO 26299.

XX Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
 KW drug; side effect; cancer; central nervous system; cardiovascular;
 KW gastrointestinal; respiratory system; single nucleotide polymorphism;
 KW SNP; cell differentiation; ds.

XX Homo sapiens.

XX WO200218632-A2.

XX 07-MAR-2002.

XX 01-SEP-2001; 2001WO-EP010074.

XX 01-SEP-2000; 2000DE-01043826.

XX 05-SEP-2000; 2000DE-01044543.

XX (EPIG-) EPIGENOMICS AG.

XX Olek A, Piepenbrock C, Berlin K, Guetig D;

XX WPI; 2002-371829/40.

PT Determining the degree of cytosine methylation in genomic DNA, useful for
 PT diagnosis and prognosis, comprises selective hybridization of amplicons
 PT from chemically treated DNA.

XX Claim 12; 56pp + Sequence Listing; 56pp; German.

XX This invention describes a novel method for determining the degree of
 CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a
 CC genomic sample of DNA. The sample is treated chemically to convert
 CC cytosine (C) but not methylated C, to uracil, then part of the genomic
 CC DNA that contains the target C is amplified to form a labeled amplicon.
 CC The amplicon is hybridised to two classes, each with at least one member,
 CC of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the
 CC degree of hybridisation to both classes is determined from the label on
 CC the amplicon. From the ratio of labels hybridised to the two classes of
 CC oligomers, the degree of methylation is calculated. The method is used:

CC (i) for diagnosis and/or prognosis of side effects of therapeutic drugs
 CC and of a wide range of diseases, e.g. cancer, disorders of the central
 CC nervous, cardiovascular, gastrointestinal and respiratory systems etc.,
 CC particularly by detecting mutations or single nucleotide polymorphisms
 CC (SNP's); and (ii) for differentiation of cell or tissue types and for
 CC investigating cell differentiation. The method allows the methylation
 CC status of many C residues to be determined simultaneously. ABQ13410-
 CC ABQ54121 represent genomic DNA sequences used to illustrate the method
 CC for determining the degree of cytosine methylation described in the
 CC disclosure of the invention

XX SQ Sequence 753 BP; 66 A; 111 C; 303 G; 273 T; 0 U; 0 Other;

Query Match 70.4%; Score 17.6; DB 6; Length 753;
 Best Local Similarity 83.3%; Pred. No. 3e+02;
 Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAACGACACAGCGCCCTCG 24
 |||||
 Db 468 AAAAAACGACCGCGACCCCGCG 445

RESULT 12
 ABQ41506/c
 ID ABQ41506 standard; DNA; 1017 BP.

XX AC ABQ41506;

XX 12-JUL-2002 (first entry)

DE Oligonucleotide for detecting cytosine methylation SEQ ID NO 28097.

XX Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
 KW drug; side effect; cancer; central nervous system; cardiovascular;
 KW gastrointestinal; respiratory system; single nucleotide polymorphism;
 KW SNP; cell differentiation; ds.

XX Homo sapiens.

XX WO200218632-A2.

XX 07-MAR-2002.

XX 01-SEP-2001; 2001WO-EP010074.

XX 01-SEP-2000; 2000DE-01043826.

XX 05-SEP-2000; 2000DE-01044543.

XX (EPIG-) EPIGENOMICS AG.

XX Olek A, Piepenbrock C, Berlin K, Guetig D;

XX WPI; 2002-371829/40.

XX Determining the degree of cytosine methylation in genomic DNA, useful for
 XX diagnosis and prognosis, comprises selective hybridization of amplicons
 XX from chemically treated DNA.

XX Claim 12; 56pp + Sequence Listing; 56pp; German.

XX This invention describes a novel method for determining the degree of
 CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a
 CC genomic sample of DNA. The sample is treated chemically to convert
 CC cytosine (C) but not methylated C, to uracil, then part of the genomic
 CC DNA that contains the target C is amplified to form a labeled amplicon.
 CC The amplicon is hybridised to two classes, each with at least one member,
 CC of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the
 CC degree of hybridisation to both classes is determined from the label on
 CC the amplicon. From the ratio of labels hybridised to the two classes of
 CC oligomers, the degree of methylation is calculated. The method is used:
 CC (i) for diagnosis and/or prognosis of side effects of therapeutic drugs
 CC and of a wide range of diseases, e.g. cancer, disorders of the central
 CC nervous, cardiovascular, gastrointestinal and respiratory systems etc.,

CC particularly by detecting mutations or single nucleotide polymorphisms
CC (SNP's); and (ii) for differentiation of cell or tissue types and for
CC investigating cell differentiation. The method allows the methylation
CC status of many C residues to be determined simultaneously. ABQ13410-
CC ABQ54121 represent genomic DNA sequences used to illustrate the method
CC for determining the degree of cytosine methylation described in the
CC disclosure of the invention

XX Sequence 1017 BP; 83 A; 148 C; 350 G; 436 T; 0 U; 0 Other;
SQ

Query Match 70.4%; Score 17.6; DB 6; Length 1017;
Best Local Similarity 83.3%; Pred. No. 3.2e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAACGACACGAGCCCTCG 24
|||||
Db 134 AAAAAACGACACGAGTACCCACG 111

RESULT 13
ABQ41507
ID ABQ41507 standard; DNA; 1017 BP.
XX
AC ABQ41507;
XX
DT 12-JUL-2002 (first entry)
XX
DE Oligonucleotide for detecting cytosine methylation SEQ ID NO 28098.
XX
KW Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
KW drug, side effect; cancer; central nervous system; cardiovascular;
KW gastrointestinal; respiratory system; single nucleotide polymorphism;
KW SNP; cell differentiation; ds.
XX
OS Homo sapiens.
XX
XX WO200218632-A2.
XX
PD 07-MAR-2002.
XX
PF 01-SEP-2001; 2001WO-EP010074.
XX
PR 01-SEP-2000; 2000DE-01043826.
XX
PR 05-SEP-2000; 2000DE-01044543.
XX
XX (EPIG-) EPIGENOMICS AG.
XX
XX Olek A, Piepenbrock C, Berlin K, Guetig D;
XX
XX WPI; 2002-371829/40.
XX
XX Determining the degree of cytosine methylation in genomic DNA, useful for
XX diagnosis and prognosis, comprises selective hybridization of amplicons
XX from chemically treated DNA.
XX
XX Claim 12; 56pp + Sequence Listing; 56pp; German.

CC This invention describes a novel method for determining the degree of
CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a
CC genomic sample of DNA. The sample is treated chemically to convert
CC cytosine (C) but not methylated C, to uracil, then part of the genomic
CC DNA that contains the target C is amplified to form a labeled amplicon.
CC The amplicon is hybridised to two classes, each with at least one member,
CC of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the
CC degree of hybridisation to both classes is determined from the label on
CC the amplicon. From the ratio of labels hybridised to the two classes of
CC oligomers, the degree of methylation is calculated. The method is used:
CC (i) for diagnosis and/or prognosis of side effects of therapeutic drugs
CC and of a wide range of diseases, e.g. cancer, disorders of the central
CC nervous, cardiovascular, gastrointestinal and respiratory systems etc.,
CC particularly by detecting mutations or single nucleotide polymorphisms
CC (SNP's); and (ii) for differentiation of cell or tissue types and for
CC investigating cell differentiation. The method allows the methylation

CC status of many C residues to be determined simultaneously. ABQ13410-
CC ABQ54121 represent genomic DNA sequences used to illustrate the method
CC for determining the degree of cytosine methylation described in the
CC disclosure of the invention

XX Sequence 1017 BP; 436 A; 350 C; 148 G; 83 T; 0 U; 0 Other;
SQ

Query Match 70.4%; Score 17.6; DB 6; Length 1017;
Best Local Similarity 83.3%; Pred. No. 3.2e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAACGACACGAGCCCTCG 24
|||||
Db 884 AAAAAACGACACGCGTACCCACG 907

RESULT 14
ABV28495
ID ABV28495 standard; cDNA; 6073 BP.
XX
XX AC ABV28495;
XX
XX 16-SEP-2002 (first entry)
XX
XX Human prostate expression marker cDNA 28486.
XX
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX
XX Homo sapiens.
XX
XX WO200160860-A2.
XX
XX 23-AUG-2001.
XX
XX 20-FEB-2001; 2001WO-US005171.
XX
XX 17-FEB-2000; 2000US-0183319P.
XX
XX 16-MAR-2000; 2000US-0189862P.
XX
XX 25-MAY-2000; 2000US-0207454P.
XX
XX 09-JUN-2000; 2000US-0211314P.
XX
XX 18-JUL-2000; 2000US-0219007P.
XX
XX 13-DEC-2000; 2000US-0255281P.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Schlegel R, Endege WO, Monahan JE;
XX
XX WPI; 2001-662795/76.
XX
XX Novel isolated nucleic acid molecule associated with cancerous state of
XX prostate cells and correlating with presence of prostate cancer, useful
XX for detecting presence of prostate cancer, stage of prostate cancer.
XX
XX Claim 1; Page 5951-5952; 11750pp; English.

CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (i) is also useful as a pharmacodynamic or pharmacogenomic marker

XX Sequence 6073 BP; 1593 A; 1457 C; 1588 G; 1428 T; 0 U; 7 Other;
SQ

Query Match 70.4%; Score 17.6; DB 5; Length 6073;
Best Local Similarity 83.3%; Pred. No. 3.9e+02;

Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAACGACACAGCGCCCTCG 24
 DB 12 AAAAAAGAGACTCGCGAGCCCTCG 35

RESULT 15

ABV22669

ID ABV22669 standard; cDNA; 6073 BP.

XX AC

XX ABV22669;

XX AC

XX 13-SEP-2002 (first entry)

XX DT

XX DE

XX Human prostate expression marker cDNA 22660.

XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynanamic marker;

KW pharmacogenomic marker; gene; ss.

KW OS

XX Homo sapiens.

XX WO200160860-A2.

XX PN

XX 23-AUG-2001.

XX PD

XX 20-FEB-2001; 2001WO-US005171.

XX PF

XX 17-FEB-2000; 2000US-0183319P.

XX PR

XX 16-MAR-2000; 2000US-0189862P.

XX PR

XX 25-MAY-2000; 2000US-0207454P.

XX PR

XX 03-JUN-2000; 2000US-0211314P.

XX PR

XX 18-JUL-2000; 2000US-0219007P.

XX PR

XX 13-DEC-2000; 2000US-0255281P.

XX XX

PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX PI

XX Schlegel R, Endege WO, Monahan JE;

XX DR

XX WPI; 2001-662795/76.

XX XX

XX Novel isolated nucleic acid molecule associated with cancerous state of

PT prostate cells and correlating with presence of prostate cancer, useful

PT for detecting presence of prostate cancer, stage of prostate cancer.

PT Claim 1; Page 3977-3978; 11750pp; English.

XX PS

XX The invention relates to an isolated nucleic acid molecule (I) comprising

CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the

CC specification or its complement. (I) is useful for: (a) assessing whether

CC a patient is afflicted with prostate cancer; (b) monitoring the efficacy

CC progression of prostate cancer in a patient; (c) assessing the efficacy

CC of a test compound to inhibit prostate cancer in a patient; (d) assessing

CC the efficacy of a therapy for inhibiting prostate cancer in a patient;

CC (e) selecting a composition for inhibiting prostate cancer in a patient;

CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)

CC determining whether prostate cancer has metastasized in a patient; (h)

CC assessing the aggressiveness or indolence of prostate cancer in a patient

CC ; (I) is also useful as a pharmacodynanamic or pharmacogenomic marker

XX SQ

Sequence 6073 BP; 1593 A; 1457 C; 1588 G; 1428 T; 0 U; 7 Other;

Query Match 70.4%; Score 17.6; DB 5; Length 6073;

Best Local Similarity 83.3%; Pred. NO. 3.9e+02;

Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAACGACACAGCGCCCTCG 24

DB 12 AAAAAAGAGACTCGCGAGCCCTCG 35

RESULT 16

ADQ97433

ID ADQ97433 standard; DNA; 295772 BP.

XX AC

XX ADQ97433;

XX 07-OCT-2004 (first entry)

XX DT

XX Human cancer associated sequence HD08-047, SEQ ID 410.

XX DE

XX Cytostatic; Gene Therapy; cancer; leukemia; lymphoma; Human; ds.

XX KW

XX Homo sapiens.

XX OS

XX WO2004060304-A2.

XX PN

XX 22-JUL-2004.

XX PD

XX 22-DEC-2003; 2003WO-US041389.

XX PF

XX 27-DEC-2002; 2002US-00330773.

XX PR

XX (SAGR-) SAGRES DISCOVERY INC.

XX PA

XX Morris DW, Malandro MS;

XX PI

XX WPI; 2004-543781/52.

XX DR

XX New isolated cancer associated nucleic acids comprising at least 10

XX PT contiguous nucleotides, useful for diagnosing, preventing and/or treating

XX PT cancers such as leukemia and lymphoma.

XX PS

XX Claim 1; SEQ ID NO 410; 199pp; English.

XX XX

XX The present invention relates to cancer associated sequences (ADQ97025-

XX ADQ98004). The sequences are useful for the diagnosis, prevention and/or

XX treatment of cancer, such as leukemia and lymphoma. Note: The sequence

XX data for this patent did not form part of the printed specification, but

XX was obtained in electronic format directly from WIPO at

XX ftp.wipo.int/pub/published_pct_sequences.

XX CC

XX Sequence 295772 BP; 74399 A; 61916 C; 68604 G; 90853 T; 0 U; 0 Other;

XX SQ

Query Match 70.4%; Score 17.6; DB 12; Length 295772;

Best Local Similarity 83.3%; Pred. No. 6.1e+02;

Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAACGACACAGCGCCCTCG 24

DB 9743 AAAAAAGAGACTCGCGAGCCCTCG 9766

RESULT 17

ADT45037

ID ADT45037 standard; cDNA; 1797 BP.

XX AC

XX ADT45037;

XX 02-DEC-2004 (first entry)

XX DT

XX Bacterial polynucleotide #19788.

XX DE

XX Recombinant DNA construct; transformed plant; improved plant property;

XX cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;

XX pathogen tolerance; pest tolerance; plant disease resistance;

XX cell cycle pathway modification; plant growth regulator;

XX homologous recombination; seed oil yield; protein yield; carbohydrate;

XX nitrogen; phosphorus; photosynthesis; lignin; galactomannan;

XX bacterial polynucleotide; gene; ss.

XX OS

XX Bacteria.

XX US2003233675-A1.

XX PN

XX 18-DEC-2003.

XX PD

XX 20-FEB-2003; 2003US-00369493.
 XX 21-FEB-2002; 2002US-0360039P.
 XX (CAOY/) CAO Y.
 PA (HINK/) HINKLE G J.
 PA (SLAT/) SLATER S C.
 PA (CHEN/) CHEN X.
 PA (GOLD/) GOLDMAN B S.
 XX
 PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
 XX WPI; 2004-061375/06.
 DR
 XX New recombinant DNA construct comprising a promoter positioned to provide
 PT for expression of a polynucleotide encoding a polypeptide from a
 PT microbial source, useful for producing plants with improved properties.
 XX
 XX Claim 1; SEQ ID NO 43475; 122pp; English.
 PS
 XX The invention relates to a recombinant DNA construct comprising a
 CC promoter functional in a plant cell, where the promoter is positioned to
 CC provide for expression of a polynucleotide encoding a polypeptide from a
 CC microbial source. The invention also relates to a transformed plant
 CC comprising the recombinant DNA construct and a method of producing a
 CC transformed plant having an improved property. The plant is a crop plant
 CC such as maize or soybean. The method of producing a transformed plant
 CC having an improved property comprises transforming a plant with the
 CC recombinant DNA construct and growing the transformed plant, where the
 CC polynucleotide or polypeptide is useful for improving plant properties.
 CC The recombinant DNA construct is useful for producing plants with
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,
 CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
 CC increased resistance to plant disease, better growth rate by modification
 CC of the cell cycle pathway with plant growth regulators, increased rate of
 CC homologous recombination, modified seed oil or protein yield and/or
 CC content, improved yield by modification of carbohydrate, nitrogen or
 CC phosphorus use and/or uptake, by modification of photosynthesis or by
 CC providing improved plant growth and development under at least one stress
 CC condition, improved lignin production or improved galactomannan
 CC production. This sequence represents a bacterial polynucleotide used in
 CC the scope of the invention. Note: The sequence data for this patent did
 CC not form part of the printed specification but was obtained in electronic
 CC format from USPTO at seqdata.uspto.gov/sequence.html.
 XX
 SQ Sequence 1797 BP; 514 A; 332 C; 408 G; 543 T; 0 U; 0 Other;
 Query Match 68.8%; Score 17.2; DB 13; Length 1797;
 Best Local Similarity 86.4%; Pred. No. 5.2e+02; Mismatches 0; Gaps 0;
 Matches 19; Conservative 0; Indels 3; Indels 0; Gaps 0;
 QY 1 AAAAAACGACACAGCGGCCT 22
 DB 336 AAAAAACGACACAGCGGCCT 357
 RESULT 18
 ACL38755
 ID ACL38755 standard; cDNA; 2000 BP.
 XX
 AC ACL38755;
 XX
 XX 02-JUN-2005 (first entry)
 DT
 DE Rice stress-regulated promoter SEQ ID NO:17318.
 XX
 KW ss; abiotic stress tolerance; transgenic plant; plant; cereal;
 XX agriculture.
 XX Oryza sativa.
 OS
 XX WO2003008540-A2.
 PN

XX 30-JAN-2003.
 PD
 XX 21-JUN-2002; 2002WO-US019668.
 PF
 XX 22-JUN-2001; 2001US-0300112P.
 PR
 XX 24-AUG-2001; 2001US-0314662P.
 PR
 XX 26-SEP-2001; 2001US-0325277P.
 PR
 XX 21-NOV-2001; 2001US-0332132P.
 XX
 PA (SYGN) SYNGENTA PARTICIPATIONS AG.
 XX
 PI Kreps J, Briggs SP, Cooper B, Glazebrook J, Goff SA, Katagiri F;
 PI Moughamer T, Provart N, Rieke D, Zhu T;
 XX WPI; 2003-248011/24.
 DR
 XX New stress-responsive nucleic acid, useful for altering the
 PT responsiveness of a plant, e.g. cereal, to an abiotic stress such as cold
 PT stress, salt stress or osmotic stress.
 XX
 XX Claim 48; SEQ ID NO 17318; 89pp; English.
 PS
 XX The invention relates to novel abiotic stress responsive polynucleotides
 CC and polypeptides. Also disclosed are vectors, expression cassettes, host
 CC cells, and plants containing such polynucleotides. Also disclosed are
 CC methods for using the polynucleotides and polypeptides to alter the
 CC responsiveness of a plant to abiotic stress. The invention is useful in
 CC agriculture. The nucleic acid is useful for determining whether a test
 CC plant has been exposed to an abiotic stress condition. It is also useful
 CC for selecting an agent that alters abiotic stress regulated
 CC polynucleotide expression in a plant cell, and to identify a homolog or
 CC ortholog to an abiotic stress responsive polynucleotide. The nucleic acid
 CC molecule and the polypeptide encoded by it are useful in altering the
 CC responsiveness of a plant to an abiotic stress, such as cold stress, salt
 CC stress, osmotic stress or any of their combinations. The present sequence
 CC is used in the exemplification of the invention
 XX
 SQ Sequence 2000 BP; 558 A; 432 C; 413 G; 597 T; 0 U; 0 Other;
 Query Match 68.8%; Score 17.2; DB 11; Length 2000;
 Best Local Similarity 86.4%; Pred. No. 5.3e+02; Mismatches 0; Gaps 0;
 Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 AAAAAACGACACAGCGGCCT 22
 DB 1636 AAAAAACGACACAGCGGCCT 1657
 RESULT 19
 ADB84162/c
 ID ADB84162 standard; DNA; 6240 BP.
 XX
 AC ADB84162;
 XX
 XX 04-DEC-2003 (first entry)
 DT
 XX
 DE Gene expression estimation method-related DNA sequence #22.
 XX
 KW gene expression; plant; flower appearance; ds.
 XX
 OS Unidentified.
 XX
 XX WO2003044227-A1.
 PN
 XX 30-MAY-2003.
 PD
 XX 21-NOV-2001; 2001WO-JP010195.
 PF
 XX 21-NOV-2001; 2001WO-JP010195.
 PR
 XX (NAAG-) NAT INST AGROBIOLOGICAL SCI.
 PA
 XX

PI Higo K, Iwamoto M;
XX WPI; 2003-449821/42.
XX
XX Identification of sequences homologous to a transposable element in genes
PT of a plant such as rice for screening genes expressed in particular
PT organs of the plant including the flower.
XX
XX Disclosure; Page 127-131; 147pp; Japanese.
XX
XX The invention comprises a method of estimating gene expression in a
CC chosen organ of a plant. The method involves searching sequences in the
CC neighbourhood of putative protein-encoding regions of genes expressed in
CC the organ which are homologous to a known transposon key sequence, and
CC then selecting for genes in which these sequences are present. The method
CC of the invention is useful for establishing the pattern of gene
CC expression in particular organs of a plant (e.g. the flower) in
CC connection with the improvement of varieties of rice and other plants and
CC modification of the structural form and appearance of flowers. The
CC present DNA sequence was used in the exemplification of the invention.
XX
XX Sequence 6240 BP; 1574 A; 1348 C; 1317 G; 2001 T; 0 U; 0 Other;
SQ
Query Match 68.8%; Score 17.2; DB 10; Length 6240;
Best Local Similarity 86.4%; Pred. No. 6e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 AAAAAACGACACGAGCCCT 22
DB 1786 AAAAAATGACACGACGACCAT 1765
RESULT 20
AAF44516/c
ID AAF44516 standard; cDNA; 283 BP.
XX
XX AAF44516;
AC
XX
XX 26-MAR-2001 (first entry)
DT
XX
XX Mouse dextran sodium sulphate induced colitis EST SEQ ID NO:46.
DE
XX
XX Mouse; gastrointestinal inflammation; DSS-induced colitis; EST;
XX dextran sodium sulphate induced colitis; expressed sequence tag;
XX inflammatory bowel disease; Crohn's disease; ulcerative colitis;
XX immunosuppressive; immunostimulant; antiarthritic; antirheumatic;
XX antianaemic; antiasthmatic; antidepressant; cerebroprotective; cardiac;
XX antiallergic; antiinflammatory; antiviral; antibacterial; antifungal;
XX vulnary; haematopoietic cell stimulator; differentiation; anaemia;
XX proliferation; haematopoietic cell; immunological deficiency syndrome;
XX blood disorder; wound healing; stroke; heart attack; asthma;
XX autoimmune disorder; rheumatoid arthritis; allergic reaction;
XX graft versus host disease; infectious disease; infection; ss.
XX
XX Mus musculus.
XX
XX WO2000077166-A2.
PN
XX
XX 21-DEC-2000.
PD
XX
XX 09-JUN-2000; 2000WO-US015973.
PP
XX
XX 10-JUN-1999; 99US-0138487P.
PR
XX
XX (DIGI-) DIGITAL GENE TECHNOLOGIES INC.
PA
XX
XX Viney J, Sims JE, Dubose RF, Hilbush BS, Hasel KW, Buchner RP;
PI
XX
XX WPI; 2001-061869/07.
DR
XX
XX New isolated nucleic acid molecule useful for prevention, treatment or
PT amelioration of a medical condition, e.g. inflammatory bowel disease.
PT
XX

PS Claim 1; Page 111; 128pp; English.
XX
XX AAF44471 to AAF44532 represent expressed sequences tags (ESTs) isolated
CC from a mouse having dextran sodium sulphate (DSS)-induced colitis (I).
CC AAB51241 represents a protein (II) encoded by (I). (I) can have
CC immunosuppressive; immunostimulant; antiarthritic; antirheumatic.
CC antianaemic; cardiac; antiasthmatic; antidepressant; cerebroprotective;
CC antiallergic; antiinflammatory; antiviral; antibacterial; antifungal and
CC vulnary activities, and are haematopoietic cell stimulators. (I) and
CC (II) can be used to prevent, treat or ameliorate a medical condition. (I)
CC and (II) can be used to increase differentiation and proliferation of
CC haematopoietic cells in the treatment of immunological deficiency
CC syndromes, blood disorders (e.g. anaemia, wound healing, stroke, heart
CC attack), autoimmune disorders (e.g. rheumatoid arthritis, asthma,
CC allergic reactions, modulate inflammation, especially inflammatory bowel
CC disease and Crohn's disease), graft versus host disease and infectious
CC diseases (e.g. viral, bacterial or fungal infection). AAF44533 to
CC AAF44598 represent PCR primers used in the exemplification of the present
CC invention
XX
XX Sequence 283 BP; 62 A; 74 C; 66 G; 81 T; 0 U; 0 Other;
SQ
Query Match 68.0%; Score 17; DB 4; Length 283;
Best Local Similarity 80.0%; Pred. No. 5.2e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 1 AAAAAACGACACGAGCCCTCGG 25
DB 54 AAAAAACGGGTCTATCGGCGCTCGG 30
RESULT 21
ABD00388/c
ID ABD00388 standard; DNA; 603 BP.
XX
XX ABD00388;
AC
XX
XX 29-JUL-2004 (first entry)
DT
XX
XX Klebsiella pneumoniae polynucleotide seqid 6163. .
DE
XX
XX Recombinant expression vector; transcription regulatory element;
XX Klebsiella pneumoniae protein; antibacterial; Vaccine; gene; ds.
XX
XX Klebsiella pneumoniae.
OS
XX
XX US6610836-B1.
PN
XX
XX 26-AUG-2003.
PD
XX
XX 27-JAN-2000; 2000US-00489039.
PF
XX
XX 29-JAN-1999; 99US-0117747P.
PR
XX
XX (GENO-) GENOME THERAPEUTICS CORP.
PA
XX
XX Breton GL, Osborne M;
PI
XX
XX WPI; 2003-895346/82.
DR
XX
XX P-PSDB; ABO66817.
DR
XX
XX New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for
PT preparing a vaccine composition against Klebsiella pneumoniae.
PT
XX
XX Disclosure; SEQ ID NO 6163; 932pp; English.
PS
XX
XX The invention describes a new isolated nucleic acid encoding a Klebsiella
CC pneumoniae polypeptide. Also described are: a recombinant expression
CC vector comprising the nucleic acid, operably linked to a transcription
CC regulatory element; and a cell comprising the recombinant expression
CC vector. The nucleic acid is useful for preparing a vaccine composition
CC against Klebsiella pneumoniae. This sequence encodes a Klebsiella
CC pneumoniae polypeptide of the invention

CC antibodies. They may be useful for treatment of autoimmune disorders
CC (multiple sclerosis, insulin dependent diabetes), allergic conditions
CC (asthma), myeloid or lymphoid cell deficiencies, wounds, burns, ulcers,
CC osteoporosis, osteoarthritis, central nervous system disorders
CC (Alzheimer's, Parkinson's, Huntington's disease, stroke), coagulation
CC disorders (haemophilia, thrombosis), inflammatory disorders (Crohn's
CC disease), tumours, bacterial, fungal or viral infections, depression and
CC psoriasis. AA445926 to AA445931 represent linker variants which are given
CC in the exemplification of the present invention
XX
SQ Sequence 746 BP; 174 A; 190 C; 181 G; 199 T; 0 U; 2 Other;

Query Match 68.0%; Score 17; DB 3; Length 746;
Best Local Similarity 80.0%; Pred. No. 5.8e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAAGACACACGAGCGCCTCG 25
||||| ||||| ||||| |||||
DB 446 AAAAACCACAATGAGCCTCAG 422
||||| ||||| ||||| |||||

RESULT 23
AAK84896
ID AAK84896 standard; DNA; 3271 BP.
XX
AC AAK84896;
XX
DT 07-NOV-2001 (first entry)
XX
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:39708.
XX
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KM cytostatic; gene therapy; vaccine; metastasis; ds.
XX
OS Homo sapiens.
XX
XN WO200157182-A2.
PN
XX
PD 09-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US0001354.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.

RESULT 24	
ADR07098	
ID	ADR07098 standard; cDNA; 4473 BP.
XX	
AC	AAS65021;
XX	
DT	13-FEB-2002 (first entry)
XX	
AC	ADR07098;
XX	
DT	04-NOV-2004 (first entry)
XX	
DE	Full length human cDNA useful for treating neurological disease Seq 604.
XX	
KW	gene; ss; human; oligo-capping method; diagnostic marker; gene therapy;
KW	osteoporosis; neurological disease; Alzheimer's disease;
KW	Parkinson's disease; dementia; short memory; cancer;
KW	sense or motor function; emotional reaction; fear response; panic;
KW	osteopathic; neuroprotective; nootropic; antiparkinsonian; cytostatic;
KW	tranquilliser.
XX	
OS	Homo sapiens.
XX	
PN	EP1447413-A2.
XX	
PD	18-AUG-2004.
XX	
PF	12-FEB-2004; 2004EP-00003145.
XX	
PR	14-FEB-2003; 2003JP-00102207.
XX	
PR	09-MAY-2003; 2003JP-00131452.
XX	
PA	(REAS-) RES ASSOC BIOTECHNOLOGY.
XX	
PI	Isozaki T, Yamamoto J, Nishikawa T, Isono Y, Sugiyama T, Otsuki T;
PI	Wakamatsu A, Ishii S, Nagai K, Irie R;
PI	WPI; 2004-583265/57.
DR	P-PSDB; ADR09054.
XX	
PT	New 1995 cDNA, useful for treating osteoporosis, neurological diseases.
PT	Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.
XX	
PS	Claim 1; SEQ ID NO 604; 2686pp; English.
XX	
CC	This invention relates to novel, isolated full length human cDNA
CC	molecules and the encoded proteins thereof. Specifically, it refers to
CC	cDNA clones obtained by an oligo-capping method, where none of these
CC	clones are identical to any known human mRNAs. The present invention
CC	describes an immunoassay to identify agonists and antagonists, as well as
CC	antibodies, antisense molecules and siRNAs that can all be used to bind
CC	to and modulate expression of the cDNA molecules. As such, these
CC	molecules are useful for diagnostic markers or therapeutic targets for
CC	the various diseases or morbid states. In particular, they are useful in
CC	gene therapy for treating osteoporosis, neurological disease, Alzheimer's
CC	disease, Parkinson's disease, dementia, short memory and various cancers,
CC	as well as for maintaining equilibrium of sense or motor function, and
CC	for treating emotional reaction, fear response and panic. Accordingly,
CC	they exhibit osteopathic, neuroprotective, nootropic, antiparkinsonian,
CC	cytostatic and tranquilliser activities. This polynucleotide is a full
CC	length human cDNA sequence of the invention. NOTE: This sequence is not
CC	given in the sequence listing of the specification but can be obtained on
CC	CD-ROM from the European Patent Office, Vienna Sub-office.
XX	
SQ	Sequence 4473 BP; 989 A; 1238 C; 1138 G; 1108 T; 0 U; 0 Other;
Query Match	68.0%; Score 17; DB 13; Length 4473;
Best Local Similarity	80.0%; Pred. No. 7.2e+02;
Matches	20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY	1 AAAAAACGACACAGCGCCCTCGG 25
DB	3640 AAAAAACACAGCCATCCCTCGG 3664
RESULT 25	
AAS65021	
ID	AAS65021 standard; cDNA; 6092 BP.
XX	
AC	AAS65021;
XX	
DT	13-FEB-2002 (first entry)
XX	
DE	DNA encoding novel human diagnostic protein #825.
XX	
KW	Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW	food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX	
OS	Homo sapiens.
XX	
PN	WO2001175067-A2.
XX	
PD	11-OCT-2001.
XX	
PF	30-MAR-2001; 2001WO-US008631.
XX	
PR	31-MAR-2000; 2000US-00540217.
XX	
PR	23-AUG-2000; 2000US-00649167.
XX	
PA	(HYSE-) HYSEQ INC.
XX	
PI	Drmanac RT, Liu C, Tang YT;
XX	
XX	WPI; 2001-639362/73.
DR	P-PSDB; ABG00834.
XX	
PT	New isolated polynucleotide and encoded polypeptides, useful in
PT	diagnostics, forensics, gene mapping, identification of mutations
PT	responsible for genetic disorders or other traits and to assess
PT	biodiversity.
XX	
PS	Claim 1; SEQ ID NO 825; 103pp; English.
XX	
CC	The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC	sequences. (I) is useful as hybridisation probes, polymerase chain
CC	reaction (PCR) primers, oligomers, and for chromosome and gene mapping.
CC	and in recombinant production of (II). The polynucleotides are also used
CC	in diagnostics as expressed sequence tags for identifying expressed
CC	genes. (I) is useful in gene therapy techniques involving (II). (II) is
CC	activity of (II) or to treat disease states involving (II). (II) is
CC	useful for generating antibodies against it, detecting or quantitating a
CC	polypeptide in tissue, as molecular weight markers and as a food
CC	supplement. (II) and its binding partners are useful in medical imaging
CC	of sites expressing (II). (I) and (II) are useful for treating disorders
CC	involving aberrant protein expression or biological activity. The
CC	polypeptide and polynucleotide sequences have applications in
CC	diagnostics, forensics, gene mapping, identification of mutations
CC	responsible for genetic disorders or other traits to assess biodiversity
CC	and to produce other types of data and products dependent on DNA and
CC	amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC	coding sequences of the invention. Note: The sequence data for this
CC	patent did not appear in the printed specification, but was obtained in
CC	electronic format directly from WIPO at
CC	ftp.wipo.int/pub/published_pct_sequences
XX	
SQ	Sequence 6092 BP; 1677 A; 1318 C; 1389 G; 1708 T; 0 U; 0 Other;
Query Match	68.0%; Score 17; DB 5; Length 6092;
Best Local Similarity	80.0%; Pred. No. 7.4e+02;
Matches	20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY	1 AAAAAACGACACAGCGCCCTCGG 25
DB	2200 AAAAAAGGAGATAGCGACCCCTCAG 2224
RESULT 26	
AEA35989	
ID	AEA35989 standard; DNA; 7318 BP.
XX	

PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241212P.
PR 20-OCT-2000; 2000US-0241285P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
(HUMA-) HUMAN GENOME SCI INC.
Rosen CA, Barash SC, Ruben SM;
PI

XX WPI; 2001-483426/52.
XX Nucleic acids encoding human immune/haematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and metastasis.
XX Disclosure; SEQ ID NO 41861; 3071pp + Sequence Listing; English.
XX AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting the
CC nucleic acids into a host cell and culturing the cell to prevent the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC represent sequences used in the exemplification of the present invention
XX SQ Sequence 16605 BP; 4009 A; 3696 C; 4212 G; 4688 T; 0 U; 0 Other;
Query Match 68.0%; Score 17; DB 4; Length 16605;
Best Local Similarity 80.0%; Pred. No. 8.4e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 1 AAAAAACGACACGCGAGCCCTCGG 25
Db 1358 AAAAAAAACACAGCCCAAGG 1382
RESULT 28
AAK65489
ID AAK65489 standard; DNA; 16605 BP.
XX AC AAK65489;
XX DT 06-NOV-2001 (first entry)
XX DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:20301.
XX KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
XX OS cytostatic; gene therapy; vaccine; metastasis; ds.
XX OS Homo sapiens.
XX PN WO200157182-A2.
XX PD 09-AUG-2001.
XX PF 17-JAN-2001; 2001WO-US001354.
XX PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-018464P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214866P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.

PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-502630/55.
XX
PT Polynucleotides encoding digestive system antigens, useful for
PT diagnosing, treating, preventing and/or prognosing disorders of the
PT digestive system, particularly cancer and cancer metastases.
XX
PS Disclosure; SEQ ID NO 3915; 986pp; English.
XX
CC The present invention provides the protein and coding sequences of a
CC number of human digestive system antigens. These can be used in the
CC diagnosis, treatment and prevention of digestive system disorders,
CC including cancer, Meckel's diverticulum, bacterial or parasitic
CC infections, appendicitis, Hirschsprung's disease, chronic colitis or
CC ulcerative colitis. The present sequence is a genomic DNA fragment
CC encoding a digestive system antigen of the invention
XX
SQ Sequence 17431 BP; 5614 A; 3722 C; 3532 G; 4563 T; 0 U; 0 Other;

Query Match 68.0%; Score 17; DB 4; Length 17431;
Best Local Similarity 80.0%; Pred. No. 8.4e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 AAAAAACGACACAGCGAGCCTCGG 25
Db 7370 AAAAAACAAAAATCGAGCCTCTG 7394

RESULT 30
AA157710
ID AAI57710 standard; DNA; 17431 BP.
XX
AC AAI57710;
XX
XX 19-OCT-2001 (first entry)
XX
DE Human colorectal cancer antigen coding sequence SEQ ID NO: 247.
XX
KW Human; colorectal cancer; colorectal cancer antigen; gene therapy; ds.
XX
OS Homo sapiens.

XX
PN WO200155350-A1.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001350.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 14-JUL-2000; 2000US-0217496P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 21-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 26-SEP-2000; 2000US-0234998P.
PR 27-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.

PR 27-SEP-2000; 2000US-0235836P.
 PR 29-SEP-2000; 2000US-0236327P.
 PR 29-SEP-2000; 2000US-0236367P.
 PR 29-SEP-2000; 2000US-0236368P.
 PR 29-SEP-2000; 2000US-0236369P.
 PR 29-SEP-2000; 2000US-0236370P.
 PR 02-OCT-2000; 2000US-0236802P.
 PR 02-OCT-2000; 2000US-0237037P.
 PR 02-OCT-2000; 2000US-0237038P.
 PR 02-OCT-2000; 2000US-0237039P.
 PR 02-OCT-2000; 2000US-0237040P.
 PR 13-OCT-2000; 2000US-0239935P.
 PR 13-OCT-2000; 2000US-0239937P.
 PR 20-OCT-2000; 2000US-0240960P.
 PR 20-OCT-2000; 2000US-0241221P.
 PR 20-OCT-2000; 2000US-0241786P.
 PR 20-OCT-2000; 2000US-0241787P.
 PR 20-OCT-2000; 2000US-0241808P.
 PR 20-OCT-2000; 2000US-0241809P.
 PR 20-OCT-2000; 2000US-0241826P.
 PR 01-NOV-2000; 2000US-0244617P.
 PR 08-NOV-2000; 2000US-0246474P.
 PR 08-NOV-2000; 2000US-0246475P.
 PR 08-NOV-2000; 2000US-0246476P.
 PR 08-NOV-2000; 2000US-0246477P.
 PR 08-NOV-2000; 2000US-0246478P.
 PR 08-NOV-2000; 2000US-0246523P.
 PR 08-NOV-2000; 2000US-0246524P.
 PR 08-NOV-2000; 2000US-0246525P.
 PR 08-NOV-2000; 2000US-0246526P.
 PR 08-NOV-2000; 2000US-0246527P.
 PR 08-NOV-2000; 2000US-0246528P.
 PR 08-NOV-2000; 2000US-0246532P.
 PR 08-NOV-2000; 2000US-0246609P.
 PR 08-NOV-2000; 2000US-0246610P.
 PR 08-NOV-2000; 2000US-0246611P.
 PR 08-NOV-2000; 2000US-0246613P.
 PR 17-NOV-2000; 2000US-0249207P.
 PR 17-NOV-2000; 2000US-0249208P.
 PR 17-NOV-2000; 2000US-0249209P.
 PR 17-NOV-2000; 2000US-0249210P.
 PR 17-NOV-2000; 2000US-0249211P.
 PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249219P.
 PR 17-NOV-2000; 2000US-0249244P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249246P.
 PR 17-NOV-2000; 2000US-0249255P.
 PR 17-NOV-2000; 2000US-0249257P.
 PR 17-NOV-2000; 2000US-0249259P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250319P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 05-DEC-2000; 2000US-0256719P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 08-DEC-2000; 2000US-0251990P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.
 (HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Barash SC, Ruben SM;
 XX WPI; 2001-457727/49.
 XX Isolated polypeptide for treating, preventing and/or prognosing
 PT disorders related to the colon and rectum including colorectal cancers
 PT and also for testing and detection e.g. diagnosis.
 XX Disclosure; SEQ ID NO 247; 522pp + Sequence Listing; English.
 XX The present invention provides the protein and coding sequences of a
 CC number of colorectal cancer antigens. These are shown in AA157547-
 CC AA157619 and AA38569-AA38641. These can be used in the diagnosis,
 CC prevention and treatment of cancer of the colon and/or rectum. The
 CC present sequence is a colorectal cancer antigen genomic sequence. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 17431 BP; 5614 A; 3722 C; 3532 G; 4563 T; 0 U; 0 Other;
 Query Match 68.0%; Score 17; DB 4; Length 17431;
 Best Local Similarity 80.0%; Pred. No. 8.4e+02;
 Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 QY 1 AAAAAACGACACGAGCGCCTCGG 25
 Db 7370 AAAAAACAAATAATGCGAGCCTCTG 7394
 RESULT 31
 ABS99887
 ID ABS99887 standard; DNA; 17431 BP.
 AC ABS99887;
 DT 18-DEC-2002 (first entry)
 XX Genomic DNA #91 encoding human colorectal cancer related protein.
 DE Human; colorectal cancer related protein; colon; rectum;
 KW colorectal cancer metastasis; gastrointestinal disorder; cytostatic;
 KW gene; ds.
 XX Homo sapiens.
 OS US2002119919-A1.
 PN 29-AUG-2002.
 PD 17-JAN-2001; 2001US-00764855.
 PF 31-JAN-2000; 2000US-0179065P.
 PR (ROSE/) ROSEN C A.
 PA (RUBE/) RUBEN S M.
 PA (BARA/) BARASH S C.
 XX Rosen CA, Ruben SM, Barash SC;
 WPI; 2002-731367/79.
 XX New colorectal cancer polypeptide for diagnosing, prognosing, preventing,
 PT and treating immune, hyperproliferative, liver, kidney, reproductive
 PT disorders and for identifying modulators of therapeutic use.
 XX Disclosure; SEQ ID NO 247; 183pp; English.
 XX The present invention relates to the isolation of novel human colorectal
 CC cancer related proteins, and polynucleotide sequences encoding them. The
 CC sequences of the invention are useful in the diagnosis, treatment,
 CC prevention and/or prognosis of the colon and/or rectum, including
 CC colorectal cancer, colorectal cancer metastases, and gastrointestinal

CC disorders such as dysphagia, peptic oesophagitis, gastric reflux,
CC irritable bowel syndrome, and peritoneal diseases. The invention also
CC describes antibodies that bind colorectal cancer related proteins,
CC vectors, host cells, and recombinant and synthetic methods for producing
CC human colorectal cancer related polynucleotides, polypeptides, and/or
CC antibodies. ABS99797-ABS99974 represent genomic sequences encoding human
CC colorectal cancer related proteins. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from the USPTO web site at
CC seqdata.uspto.gov/psipdDIDentry.html
XX

SQ Sequence 17431 BP; 5614 A; 3722 C; 3532 G; 4563 T; 0 U; 0 Other;

Query Match 68.0%; Score 17; DB 6; Length 17431;

Best Local Similarity 80.0%; Pred. No. 8.4e+02;

Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAAACGACACAGCGAGCCCTCGG 25

||||||| | |||||

Db 7370 AAAAAACAAAATGCGAGCCCTCTG 7394

RESULT 32

ADB93040

ID ADB93040 standard; DNA; 17431 BP.

XX

AC ADB93040;

XX

DT 04-DEC-2003 (first entry)

XX

DE Human colorectal cancer related polypeptide DNA #91.

XX

KW ds; gene; human; colorectal cancer; antigen; gene therapy;

KW gastrointestinal disorder; inflammatory disease; infection; cancer;

KW intestinal neoplasm; small intestine carcinoma; tumor;

KW small intestine non-Hodgkin's lymphoma; small bowel lymphoma; ulcer;

KW peptic ulcer; Bruton's disease; X linked infantile agammaglobulinemia;

KW severe combined immunodeficiency; DiGeorge anomaly;

KW hyperproliferative disorder; acute lymphoblastic leukaemia;

KW acute lymphocytic leukaemia; urinary system disorder; cortical necrosis;

KW kidney infarction; cardiovascular disorder; carcinoma heart disease;

KW arrhythmia; respiratory disorder; non-allergic rhinitis; sinusitis;

KW musculoskeletal system disorder; Albers-Schonberg disease;

KW Marfan's syndrome; neurological disease; phenylketonuria;

KW Wernicke's encephalopathy; Alzheimer's disease; endocrine disorder;

KW Grave's disease; Cushing's syndrome; reproductive system disorder;

KW prostatitis; benign prostatic hypertrophy; benign prostatic hyperplasia;

KW thrombosis; atherosclerosis; myocardial infarction; ischaemic attack.

XX Homo sapiens.

XX

PN US2003054420-A1.

XX

PD 20-MAR-2003.

XX

PF 11-FEB-2002; 2002US-00072349.

XX

PR 31-JAN-2000; 2000US-0179065P.

PR 04-FEB-2000; 2000US-0180628P.

PR 24-FEB-2000; 2000US-0184664P.

PR 02-MAR-2000; 2000US-0186350P.

PR 16-MAR-2000; 2000US-0189874P.

PR 17-MAR-2000; 2000US-0190076P.

PR 18-APR-2000; 2000US-0198123P.

PR 19-MAY-2000; 2000US-0205515P.

PR 07-JUN-2000; 2000US-0209467P.

PR 28-JUN-2000; 2000US-0214886P.

PR 30-JUN-2000; 2000US-0215135P.

PR 07-JUL-2000; 2000US-0216647P.

PR 11-JUL-2000; 2000US-0216880P.

PR 11-JUL-2000; 2000US-0217487P.

PR 11-JUL-2000; 2000US-0217496P.

PR 14-JUL-2000; 2000US-0218290P.

PR 26-JUL-2000; 2000US-0220963P.

PR 26-JUL-2000; 2000US-0220964P.

PR 14-AUG-2000; 2000US-0224518P.

PR 14-AUG-2000; 2000US-0224519P.

PR 14-AUG-2000; 2000US-0225213P.

PR 14-AUG-2000; 2000US-0225214P.

PR 14-AUG-2000; 2000US-0225266P.

PR 14-AUG-2000; 2000US-0225267P.

PR 14-AUG-2000; 2000US-0225268P.

PR 14-AUG-2000; 2000US-0225270P.

PR 14-AUG-2000; 2000US-0225447P.

PR 14-AUG-2000; 2000US-0225757P.

PR 14-AUG-2000; 2000US-0225758P.

PR 14-AUG-2000; 2000US-0225759P.

PR 18-AUG-2000; 2000US-0226279P.

PR 22-AUG-2000; 2000US-0226681P.

PR 22-AUG-2000; 2000US-0226868P.

PR 22-AUG-2000; 2000US-0227182P.

PR 23-AUG-2000; 2000US-0227009P.

PR 30-AUG-2000; 2000US-0228924P.

PR 01-SEP-2000; 2000US-0229287P.

PR 01-SEP-2000; 2000US-0229343P.

PR 01-SEP-2000; 2000US-0229344P.

PR 01-SEP-2000; 2000US-0229345P.

PR 05-SEP-2000; 2000US-0229509P.

PR 05-SEP-2000; 2000US-0229513P.

PR 06-SEP-2000; 2000US-0230437P.

PR 06-SEP-2000; 2000US-0230438P.

PR 08-SEP-2000; 2000US-0231242P.

PR 08-SEP-2000; 2000US-0231243P.

PR 08-SEP-2000; 2000US-0231244P.

PR 08-SEP-2000; 2000US-0231413P.

PR 08-SEP-2000; 2000US-0231414P.

PR 08-SEP-2000; 2000US-0232080P.

PR 08-SEP-2000; 2000US-0232081P.

PR 12-SEP-2000; 2000US-0231968P.

PR 14-SEP-2000; 2000US-0232397P.

PR 14-SEP-2000; 2000US-0232398P.

PR 14-SEP-2000; 2000US-0232399P.

PR 14-SEP-2000; 2000US-0232400P.

PR 14-SEP-2000; 2000US-0232401P.

PR 14-SEP-2000; 2000US-0233063P.

PR 14-SEP-2000; 2000US-0233064P.

PR 14-SEP-2000; 2000US-0233065P.

PR 21-SEP-2000; 2000US-0234223P.

PR 21-SEP-2000; 2000US-0234274P.

PR 25-SEP-2000; 2000US-0234997P.

PR 25-SEP-2000; 2000US-0234998P.

PR 26-SEP-2000; 2000US-0235484P.

PR 27-SEP-2000; 2000US-0235834P.

PR 27-SEP-2000; 2000US-0235836P.

PR 29-SEP-2000; 2000US-0236327P.

PR 29-SEP-2000; 2000US-0236367P.

PR 29-SEP-2000; 2000US-0236368P.

PR 29-SEP-2000; 2000US-0236369P.

PR 29-SEP-2000; 2000US-0236370P.

PR 02-OCT-2000; 2000US-0236802P.

PR 02-OCT-2000; 2000US-0237037P.

PR 02-OCT-2000; 2000US-0237038P.

PR 02-OCT-2000; 2000US-0237039P.

PR 02-OCT-2000; 2000US-0237040P.

PR 13-OCT-2000; 2000US-0239315P.

PR 20-OCT-2000; 2000US-0240960P.

PR 20-OCT-2000; 2000US-0241221P.

PR 20-OCT-2000; 2000US-0241785P.

PR 20-OCT-2000; 2000US-0241786P.

PR 20-OCT-2000; 2000US-0241787P.

PR 20-OCT-2000; 2000US-0241808P.

PR 20-OCT-2000; 2000US-0241809P.

PR 20-OCT-2000; 2000US-0241826P.

PR 01-NOV-2000; 2000US-0244617P.

PR 08-NOV-2000; 2000US-0246474P.

PR 08-NOV-2000; 2000US-0246475P.

DT 18-NOV-2004 (first entry)
XX Murine cancer-associated (CA) gene MD07-126.
DE
XX
KW Mouse; cancer-associated protein; CAP; cancer-associated gene; CA; gene;
KW db; cancer; cytostatic.
XX
XX
OS Mus musculus.
XX
PN WO2004058146-A2.
XX
PD 15-JUL-2004.
XX
XX
PF 15-DEC-2003; 2003WO-US040081.
XX
XX
PR 17-DEC-2002; 2002US-00322281.
XX
XX (SAGR-) SAGRES DISCOVERY INC.
XX
XX Morris DW, Malandro MS;
PI
XX WPI; 2004-499109/47.
DR
XX Novel human cancer associated protein encoded within open reading frame
PT of cancer associated gene, useful as targets for diagnosing cancer.
PT
XX
XX Disclosure; SEQ ID NO 849; 182pp; English.
PS
XX The invention relates to cancer-associated proteins (CAP) and the cancer-
CC associated (CA) nucleic acids encoding them. The invention also relates
CC to a method for treating cancers involving administering to a patient an
CC inhibitor of CAP, and a method of screening for anticancer activity in a
CC potential drug involving providing a cell that expresses a CA gene,
CC contacting a tissue sample derived from a cancer cell with an anticancer
CC drug candidate and monitoring the effect of the anticancer drug candidate
CC on expression of the CA gene. The CAP proteins are useful for detecting
CC cancer associated with expression of a CAP protein in a test cell sample
CC and for screening for a bioactive agent capable of modulating the
CC activity of a CAP protein. The CA nucleic acids are useful for diagnosing
CC cancer, involving determining the expression of a CA nucleic acid in a
CC tissue. This sequence represents a murine CA gene of the invention. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 62278 BP; 19201 A; 12593 C; 12586 G; 16957 T; 0 U; 941 Other;
Query Match 68.0%; Score 17; DB 13; Length 62278;
Best Local Similarity 80.0%; Pred. No. 9.8e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Qy 1 AAAAAACGACACAGCGAGCCCTCGG 25
Db 28443 AAAAAACGGTCAATCGGCGCCCTCGG 28467
RESULT 35
ADD14752/c
ID ADD14752 standard; cDNA; 86765 BP.
XX
AC ADD14752;
XX
XX 01-JAN-2004 (first entry)
XX Human src biomarker polynucleotide SEQ ID NO:146.
DE
XX predictor set; protein tyrosine kinase activity modulator;
KW protein tyrosine kinase pathway; protein tyrosine kinase; cytostatic;
KW gene therapy; drug sensitivity; genetic profile; cancer; human; gene; ss.
XX
XX Homo sapiens.
OS
XX WO2003062395-A2.
PN

XX 31-JUL-2003.
PD
XX 17-JAN-2003; 2003WO-US001981.
PF
XX 18-JAN-2002; 2002US-0350061P.
PR
XX (BRIM) BRISTOL-MYERS SQUIBB CO.
PA
XX Huang F, Fairchild CR, Lee FY, Shaw P;
PI WPI; 2003-636735/60.
XX P-PSDB; ADD14150.
DR
XX New polynucleotides and polypeptides for predicting the activity of
PT compounds that interact with protein tyrosine kinases and/or protein
PT tyrosine kinase pathways.
PT
XX Claim 2; SEQ ID NO 146; 139pp; English.
PS
XX The present invention describes a predictor set comprising a plurality of
XX polynucleotides or polypeptides whose expression pattern is predictive of
CC the response of cells to treatment with a compound that modulates protein
CC tyrosine kinase activity or members of the protein tyrosine kinase
CC pathway. Also described: (1) predicting whether a compound is capable of
CC modulating the activity of cells, comprising obtaining a sample of cells,
CC determining whether the cells express a plurality of markers, and
CC correlating the expression of the markers to the compound's ability to
CC modulate the activity of the cells; (2) a plurality of cell lines for
CC identifying polynucleotides and polypeptides whose expression levels
CC correlate with compound sensitivity or resistance of cells associated
CC with a disease state; and (3) identifying polynucleotides and
CC polypeptides that predict compound sensitivity or resistance of cells
CC associated with a disease state, comprising subjecting the plurality of
CC cell lines to one or more compounds, analysing the expression pattern of
CC a microarray of polynucleotides or polypeptides, and selecting
CC polynucleotides or polypeptides that predict the sensitivity or
CC resistance of cells associated with a disease state by using the
CC expression pattern of the microarray. The polynucleotides and
CC polypeptides have cytostatic activities, and can be used in gene therapy.
CC The polynucleotides and polypeptides are useful in predicting the
CC activity of compounds that interact with protein tyrosine kinases and/or
CC protein tyrosine kinase pathways. These may be used in determining drug
CC sensitivity in patients to allow the development of individualized
CC genetic profiles which aid in treating diseases and disorders (e.g.
CC cancer) based on patient response at a molecular level. The present
CC sequence is used in the exemplification of the present invention.
XX
SQ Sequence 86765 BP; 24568 A; 21025 C; 20504 G; 20568 T; 0 U; 0 Other;
Query Match 68.0%; Score 17; DB 10; Length 86765;
Best Local Similarity 80.0%; Pred. No. 1e+03;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Qy 1 AAAAAACGACACAGCGAGCCCTCGG 25
Db 51218 AAAAAAACACAGCCGCCAAGG 51194
RESULT 36
ACN44296
ID ACN44296 standard; DNA; 115284 BP.
XX
AC ACN44296;
XX
XX 18-NOV-2004 (first entry)
DT
XX Mouse genomic sequence MCG17138.
DE
XX Cytostatic; carcinoma; lymphoma; cancer; murine; gene; ss.
KW
XX Mus musculus.
OS
XX

CC binds to a polypeptide, which is prepared by immunizing a host animal
CC with a composition comprising the polypeptide or its antigen binding
CC fragment and collecting cells from the host expressing antibodies against
CC the antigen or its antigen binding fragment, a composition comprising the
CC antibody and a carrier, a method of screening for anticancer activity, a
CC method of detecting a CA nucleic acid, a method of diagnosing cancer, a
CC method of treating cancer and a method of inhibiting expression of a CA
CC nucleic acid in a cell. The CA nucleic acids are useful for detecting CA
CC nucleic acids. The antibody is useful for detecting the presence or
CC absence of cancer cells in an individual which involves contacting cells
CC from the individual with the antibody and detecting a complex of a CA
CC protein from the cancer cells and the antibody, where the detection of
CC the complex correlates with the presence of cancer cells in the
CC individual. The composition is useful for inhibiting growth of cancer
CC cells in an individual or for delivering a therapeutic agent to cancer
CC cells in an individual. The invention is also useful for diagnosing
CC cancer, for treating cancer and for inhibiting expression of a CA gene in
CC a cell. This sequence represents murine cancer-associated genomic DNA of
CC the invention.

XX Sequence 163701 BP; 42438 A; 35611 C; 35831 G; 45904 T; 0 U; 3917 Other;
SQ

Query Match 68.0%; Score 17; DB 14; Length 163701;
Best Local Similarity 80.0%; Pred. No. 1.1e+03;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 AAAAAAGCAGCAGCGCCCTCGG 25

Db 113080 AAAAAAGCAGCAGCGCCCTCGG 113056

RESULT 39

ABK83564/c

ID ABK83564 standard; cDNA; 198161 BP.

XX AC ABK83564;

XX 14-AUG-2002 (first entry)

XX Human cDNA differentially expressed in granulocytic cells #135.

XX Human; ss; granulocytic cell; DNA chip; bacterial infection;
XX viral infection; parasitic infection; protozoal infection;
XX fungal infection; sterile inflammatory disease; psoriasis;
XX rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;
XX cardiac reperfusion injury; renal reperfusion injury; ARDS;
XX adult respiratory distress syndrome; inflammatory bowel disease;
XX Crohn's disease; ulcerative colitis; periodontal disease;
XX granulocyte activation; chronic inflammation; allergy.

XX Homo sapiens.

XX WO200228999-A2.

XX 11-APR-2002.

XX 03-OCT-2001; 2001WO-US030821.

XX 03-OCT-2000; 2000US-0237189P.

XX (GENE-) GENE LOGIC INC.

XX Beazer-Barclay Y, Weissman SM, Yamaga S, Vockley J;

XX WPI; 2002-435328/46.

XX Detecting granulocyte activation by detecting differential expression of
XX genes associated with granulocyte activation, which serves as diagnostic
XX markers that is useful for monitoring disease states and drug toxicity.

XX Claim 1; SEQ ID NO 135; 114pp; English.

XX The invention relates to detecting (M1) granulocyte (GC) activation

CC (GCA), by detecting the level of expression of gene(s) (Gs) identified by
CC DNA chip analysis as given in the specification, and comparing the
CC expression level to an expression level in an unactivated GC, where
CC differential expression of Gs is indicative of GCA. Also included are
CC modulating (M2) GA by contacting GC with an agent that alters the
CC expression of at least one gene in Gs; (2) screening (M3) for an agent
CC capable of modulating GCA or an inflammation (especially chronic) in a
CC tissue, an allergic response in a subject, exposure of a subject to a
CC pathogen or sterile inflammatory disease using the gene expression
CC profile; (3) detecting (M4) an inflammation (especially chronic) in a
CC tissue, an allergic response in a subject, exposure of a subject to a
CC pathogen or sterile inflammatory disease, by detecting the level of
CC expression in a sample of the tissue of gene(s) from Gs, where the level
CC of expression of the gene is indicative of inflammation; (4) treating
CC (M5) an inflammation (especially chronic) or in a tissue, an allergic
CC response in a subject, exposure of a subject to a pathogen or sterile
CC inflammatory disease, by contacting a tissue having inflammation with an
CC agent that modulates the expression of gene(s) from Gs in the tissue. M1
CC is useful for detecting GCA; M2 is useful for modulating GA; M3 is useful
CC for screening an agent capable of modulating GCA preferably in an
CC inflammation in a tissue; M4 is useful for detecting an inflammation
CC (especially chronic) in a tissue, an allergic response in a subject,
CC exposure of a subject to a pathogen or sterile inflammatory disease (e.g.
CC psoriasis, rheumatoid arthritis, glomerulonephritis, asthma, thrombosis,
CC cardiac reperfusion injury, renal reperfusion injury, ARDS, adult
CC respiratory distress syndrome, inflammatory bowel disease, Crohn's
CC disease, ulcerative colitis, periodontal disease; also bacterial
CC infection, viral infection, parasitic infection, protozoal infection,
CC fungal infection and M5 is useful for treating one of the above
CC conditions. The present sequence represents a gene differentially
CC expressed in granulocytes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 198161 BP; 61561 A; 45946 C; 41360 G; 49294 T; 0 U; 0 Other;

Query Match 68.0%; Score 17; DB 6; Length 198161;
Best Local Similarity 80.0%; Pred. No. 1.1e+03;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 AAAAAAGCAGCAGCGCCCTCGG 25

Db 78717 AAAAAAGCAGCGCTTGCCTAGG 78693

RESULT 40

ADQ17348/c

ID ADQ17348 standard; DNA; 198161 BP.

XX AC ADQ17348;

XX 26-AUG-2004 (first entry)

XX Human soft tissue sarcoma-upregulated DNA - SEQ ID 165.

XX soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human;
XX ds.

XX Homo sapiens.

XX WO2004048938-A2.

XX 10-JUN-2004.

XX 26-NOV-2003; 2003WO-US038193.

XX 26-NOV-2002; 2002US-0429739P.

XX (PROT-) PROTEIN DESIGN LABS INC.

XX Aziz N, Ginsburg WM, Zlotnik A;

XX

DR WPI; 2004-441208/41.
XX
PT Early detection of soft tissue sarcoma comprises determining expression
PT of a gene in a first soft tissue sample and a normal soft tissue sample
PT and comparing the gene expression, also useful in treating soft tissue
PT sarcoma.
XX
PS Example 2; SEQ ID NO 165; 210pp; English.
XX
CC The invention relates to a novel method for detecting soft tissue sarcoma
CC which comprises obtaining a first soft tissue sample from an individual
CC and a normal soft tissue sample from the same or different individual,
CC determining the expression of a gene in both samples and comparing the
CC expression of the gene in both soft tissue samples, where a higher level
CC of protein expression in the first soft tissue sample indicates the
CC presence of soft tissue sarcoma. The method of the invention has
CC cytostatic applications and may be useful for detecting soft tissue
CC sarcoma, possibly via gene therapy or vaccine production. The nucleic
CC acid sequences may be useful in diagnostic and screening applications.
CC The current sequence is that of a human soft tissue sarcoma-upregulated
CC DNA of the invention. The current sequence is not shown within the
CC specification per se but was submitted in CD format by the inventor.
XX
SQ Sequence 198161 BP; 61561 A; 45946 C; 41360 G; 49294 T; 0 U; 0 Other;
Query Match 68.0%; Score 17; DB 12; Length 198161;
Best Local Similarity 80.0%; Pred. No. 1.1e+03;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 1 AAAAAACGACACAGCGAGCCCTCGG 25
Db 78717 AAAAAAGACCAGCTTGCCCTAGG 78693
RESULT 41
ADR52701/c
ID ADR52701 standard; DNA; 198161 BP.
XX
AC ADR52701;
XX
DT 18-NOV-2004 (first entry)
XX
DE Drug therapy altered expressed gene #52.
XX
KW drug activity monitoring; expression profile; gene expression;
KW peripheral blood sample; peripheral blood mononuclear cell; drug therapy;
KW CCI-779; immunosuppressant; rapamycin; mammalian target of rapamycin;
KW mTOR; ds.
XX
OS Homo sapiens.
XX
PN WO2004072265-A2.
XX
PD 26-AUG-2004.
XX
PF 11-FEB-2004; 2004WO-US004118.
XX
PR 11-FEB-2003; 2003US-0446133P.
PR 03-APR-2003; 2003US-0459782P.
PR 23-JAN-2004; 2004US-0538246P.
XX
PA (AMHP) WYETH.
PA (BURC/) BURCZYNSKI M.
PA (TWIN/) TWINE N.
PA (DORN/) DORNER A J.
PA (TREP/) TREPICCHIO W L.
XX
PI Burczynski M, Twine N, Dornier AJ, Trepicchio WL;
XX
DR WPI; 2004-642301/62.
XX
PT Monitoring drug activities in vivo comprises comparing an expression
PT profile of a gene in a peripheral blood sample of a patient before and

PT after drug therapy.
XX
PS Disclosure; SEQ ID NO 52; 136pp; English.
XX
CC The invention relates to a method of monitoring drug activities in vivo
CC by comparing an expression profile of at least one gene in a peripheral
CC blood sample of a patient to a reference expression profile of the at
CC least one gene, where the at least one gene is differentially expressed
CC in peripheral blood mononuclear cells (PBMCs) of patients who have a non-
CC blood disease and are subjected to a drug therapy as compared to PBMCs
CC isolated from the patient before the drug therapy, and where the patient
CC has the non-blood disease and is being treated by the drug therapy. The
CC method, kit, and nucleic acid array are useful for monitoring drug
CC activities in vivo. The drug is especially CCI-779, an ester analogue of
CC the immunosuppressant rapamycin which is a potent inhibitor of the
CC mammalian target of rapamycin (mTOR). This sequence represents a gene
CC expressed in PBMC altered by the drug therapy. (Note: this sequence does
CC no form part of the printed specification but was obtained in electronic
CC format from WIPO at ftp.wipo.int/pub/published_pct_sequences/).
XX
SQ Sequence 198161 BP; 61561 A; 45946 C; 41360 G; 49294 T; 0 U; 0 Other;
Query Match 68.0%; Score 17; DB 13; Length 198161;
Best Local Similarity 80.0%; Pred. No. 1.1e+03;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 1 AAAAAACGACACAGCGAGCCCTCGG 25
Db 78717 AAAAAAGACCAGCTTGCCCTAGG 78693
RESULT 42
ADQ97430
ID ADQ97430 standard; DNA; 208765 BP.
XX
AC ADQ97430;
XX
DT 07-OCT-2004 (first entry)
XX
DE Mouse cancer associated sequence MD08-047, SEQ ID 407.
XX
KW Cytostatic; Gene Therapy; cancer; leukemia; lymphoma; Mouse; ds.
XX
OS Mus musculus.
XX
PN WO2004060304-A2.
XX
PD 22-JUL-2004.
XX
PF 22-DEC-2003; 2003WO-US041389.
XX
PR 27-DEC-2002; 2002US-00330773.
XX
PA (SAGR-) SAGRES DISCOVERY INC.
XX
PI Morris DW, Malandro MS;
XX
DR WPI; 2004-543781/52.
XX
PT New isolated cancer associated nucleic acids comprising at least 10
PT contiguous nucleotides, useful for diagnosing, preventing and/or treating
PT cancers such as leukemia and lymphoma.
XX
PS Claim 1; SEQ ID NO 407; 199pp; English.
XX
CC The present invention relates to cancer associated sequences (ADQ97025-
CC ADQ98004). The sequences are useful for the diagnosis, prevention and/or
CC treatment of cancer, such as leukemia and lymphoma. Note: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 208765 BP; 53491 A; 42141 C; 46142 G; 66771 T; 0 U; 220 Other;

db. Homo sapiens. WO2004048938-A2. 10-JUN-2004. 26-NOV-2003; 2003WO-US038193. 26-NOV-2002; 2002US-0429739P. (PROT-) PROTEIN DESIGN LABS INC. Aziz N, Ginsburg WM, Zlotnik A; WPI; 2004-441208/41. Early detection of soft tissue sarcoma comprises determining expression of a gene in a first soft tissue sample and a normal soft tissue sample and comparing the gene expression, also useful in treating soft tissue sarcoma. Example 2; SEQ ID NO 6042; 210pp; English. The invention relates to a novel method for detecting soft tissue sarcoma which comprises obtaining a first soft tissue sample from an individual, and a normal soft tissue sample from the same or different individual, determining the expression of a gene in both samples and comparing the expression of the gene in both soft tissue samples, where a higher level of protein expression in the first soft tissue sample indicates the presence of soft tissue sarcoma. The method of the invention has cytostatic applications and may be useful for detecting soft tissue sarcoma, possibly via gene therapy or vaccine production. The nucleic acid sequences may be useful in diagnostic and screening applications. The current sequence is that of a human soft tissue sarcoma-upregulated DNA of the invention. The current sequence is not shown within the specification per se but was submitted in CD format by the inventor. Sequence 2217 BP; 504 A; 592 C; 699 G; 422 T; 0 U; 0 Other; Query Match 67.2%; Score 16.8; DB 12; Length 2217; Best Local Similarity 90.0%; Pred. No. 8.2e+02; Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0; Qy 4 AAACGACACAGCGAGCCCTC 23 Db 163 AAACGACACAGCGAGCCCTC 182 RESULT 45 ABV14306 ID ABV14306 standard; cDNA; 354 BP. AC ABV14306; XX 13-SEP-2002 (first entry) XX Human prostate expression marker cDNA 14297. DE Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker; KW pharmacogenomic marker; gene; ss. XX Homo sapiens. OS WO200160860-A2. FN 23-AUG-2001. XX 20-FEB-2001; 2001WO-US005171. XX 17-FEB-2000; 2000US-0183319P. PR 16-MAR-2000; 2000US-0189862P.

```
PR 25-MAY-2000; 2000US-0207454P.
PR 09-JUN-2000; 2000US-0211314P.
PR 18-JUL-2000; 2000US-0219007P.
PR 13-DEC-2000; 2000US-0255281P.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
PA
PA Schlegel R, Endege WO, Monahan JE;
XX
XX WPI; 2001-662795/76.
XX
XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX
XX Claim 1; Page 2387; 11750pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
XX Sequence 354 BP; 99 A; 92 C; 78 G; 85 T; 0 U; 0 Other;
SQ
Query Match 66.4%; Score 16.6; DB 5; Length 354;
Best Local Similarity 82.6%; Pred. No. 8.2e+02;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 AAAAAACGACACAGCGAGCCCTC 23
Db 44 AAAAAAGACACAGCGAGCTCCC 66
RESULT 46
ABV05137
ID ABV05137 standard; cDNA; 399 BP.
XX
XX ABV05137;
AC
XX
XX 13-SEP-2002 (first entry)
DT
XX
XX Human prostate expression marker cDNA 5128.
DE
XX
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX
XX Homo sapiens.
OS
XX
XX WO200160860-A2.
PN
XX
XX 23-AUG-2001.
PD
XX
XX 20-FEB-2001; 2001WO-US005171.
PF
XX
XX 17-FEB-2000; 2000US-0183319P.
PR
XX 16-MAR-2000; 2000US-0189862P.
PR
XX 25-MAY-2000; 2000US-0207454P.
PR
XX 09-JUN-2000; 2000US-0211314P.
PR
XX 18-JUL-2000; 2000US-0219007P.
PR
XX 13-DEC-2000; 2000US-0255281P.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
PA
PA Schlegel R, Endege WO, Monahan JE;
XX
XX WPI; 2001-662795/76.
XX
XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX
XX Claim 1; Page 2387; 11750pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
XX Sequence 354 BP; 99 A; 92 C; 78 G; 85 T; 0 U; 0 Other;
SQ
Query Match 66.4%; Score 16.6; DB 5; Length 354;
Best Local Similarity 82.6%; Pred. No. 8.2e+02;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 AAAAAACGACACAGCGAGCCCTC 23
Db 44 AAAAAAGACACAGCGAGCTCCC 66
RESULT 47
ADQ20479
ID ADQ20479 standard; DNA; 401 BP.
XX
XX ADQ20479;
AC
XX
XX 26-AUG-2004 (first entry)
DT
XX
XX Human soft tissue sarcoma-upregulated DNA - SEQ ID 3299.
DE
XX
XX soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human;
KW ds.
XX
XX Homo sapiens.
OS
XX
XX WO2004048938-A2.
PN
XX
XX 10-JUN-2004.
PD
XX
XX 26-NOV-2003; 2003WO-US038193.
PF
XX
XX 26-NOV-2002; 2002US-0429739P.
PR
XX
XX (PROT-) PROTEIN DESIGN LABS INC.
PA
XX
XX Aziz N, Ginsburg WM, Zlotnik A;
PI
XX
XX WPI; 2004-441208/41.
XX
XX Early detection of soft tissue sarcoma comprises determining expression
PT of a gene in a first soft tissue sample and a normal soft tissue sample
PT and comparing the gene expression, also useful in treating soft tissue
PT sarcoma.
XX
XX Example 2; SEQ ID NO 3299; 210pp; English.
PS
XX
XX The invention relates to a novel method for detecting soft tissue sarcoma
CC which comprises obtaining a first soft tissue sample from an individual,
CC and a normal soft tissue sample from the same or different individual,
CC determining the expression of a gene in both samples and comparing the
CC expression of the gene in both soft tissue samples, where a higher level
```

CC of protein expression in the first soft tissue sample indicates the
CC presence of soft tissue sarcoma. The method of the invention has
CC cytotatic applications and may be useful for detecting soft tissue
CC sarcoma, possibly via gene therapy or vaccine production. The nucleic
CC acid sequences may be useful in diagnostic and screening applications.
CC The current sequence is that of a human soft tissue sarcoma-upregulated
CC DNA of the invention. The current sequence is not shown within the
CC specification per se but was submitted in CD format by the inventor.

XX SQ Sequence 401 BP; 98 A; 79 C; 95 G; 111 T; 0 U; 18 Other;
SQ Sequence 401 BP; 98 A; 79 C; 95 G; 111 T; 0 U; 18 Other;
Query Match 66.4%; Score 16.6; DB 12; Length 401;
Best Local Similarity 79.2%; Pred. No. 8.3e+02;
Matches 19; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 1 AAAAAACGACACAGCGCCCTCG 24
Db 102 AAAAAACTACAAAGGANCCTCG 125

RESULT 48
ACF85290
ID ACF85290 standard; DNA; 412 BP.
XX AC ACF85290;
XX DT 02-JUN-2005 (first entry)
XX DE Human SIRS/sepsis diagnostic marker DNA fragment 4150.
XX KW Systemic inflammatory response syndrome; SIRS; antibacterial;
XX KW immunosuppressive; antiinflammatory; diagnosis; sepsis; ds.
XX OS Homo sapiens.
XX PN WO2004087949-A2.
XX PD 14-OCT-2004.
XX PF 31-MAR-2004; 2004WO-EP003419.
XX PR 02-APR-2003; 2003DE-01015031.
XX PR 08-AUG-2003; 2003DE-01036511.
XX PR 02-SEP-2003; 2003DE-01040395.
XX PA (SIRS-) SIRS LAB GMBH.
XX PI Russwurm S, Reinhart K, Saluz H, Straube E, Zipfel PF, Deigner H;
XX PF 2004-748070/73.
XX DR WPI; 2004-748070/73.
XX PT In vitro detection of systemic inflammatory response syndrome and related
XX PT conditions, for e.g. monitoring progression, comprises detecting abnormal
XX PT expression of disease-related genes.
XX PS Disclosure; Page; 75pp; German.
XX CC The invention relates to a novel method for in vitro detection of
XX CC systemic inflammatory response syndrome (SIRS). The method comprises
XX CC detecting abnormal expression of disease-related genes, or their
XX CC associated peptides. The method of the invention demonstrates
XX CC antibacterial, immunosuppressive and antiinflammatory applications and
XX CC may be used for early differential diagnosis, monitoring progression, and
XX CC assessing risk, assessing the likely response to treatment and for post
XX CC mortem diagnosis of systemic inflammatory response syndrome, sepsis and
XX CC sepsis-like conditions. The recombinant or synthetic nucleic acid
XX CC sequences of the invention, or derived proteins or peptides, may be
XX CC useful as calibrants in assays for the specified diseases, for evaluating
XX CC activity or toxicity in screening for active agents and/or for
XX CC preparation of agents for treatment or prevention of the specified
XX CC diseases. The current sequence is that of a human SIRS/sepsis diagnostic
XX CC marker DNA fragment of the invention. Note: The sequence data for this
XX CC patent did not form part of the printed specification, but was obtained

CC in electronic format directly from WIPO at ftp.wipo.int/pub/published
CC pct sequences. Furthermore, a number of arbitrary SEQ ID NO.s are
CC disclosed within the specification, however, these have not been taken
CC into account during indexing due to inconsistencies in application and
CC format
XX SQ Sequence 412 BP; 138 A; 61 C; 65 G; 148 T; 0 U; 0 Other;
SQ Sequence 412 BP; 138 A; 61 C; 65 G; 148 T; 0 U; 0 Other;
Query Match 66.4%; Score 16.6; DB 13; Length 412;
Best Local Similarity 82.6%; Pred. No. 8.3e+02;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 AAAAAACGACACAGCGCCCTC 23
Db 192 AAAAAATGATAGAGGCGCCCTC 214

RESULT 49
ABV44224
ID ABV44224 standard; cDNA; 459 BP.
XX AC ABV44224;
XX DT 16-SEP-2002 (first entry)
XX DE Human prostate expression marker cDNA 44215.
XX KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX KW pharmacogenomic marker; gene; ss.
XX OS Homo sapiens.
XX PN WO200160860-A2.
XX PD 23-AUG-2001.
XX PF 20-FEB-2001; 2001WO-US005171.
XX PR 17-FEB-2000; 2000US-0183319P.
XX PR 16-MAR-2000; 2000US-0189862P.
XX PR 25-MAY-2000; 2000US-0207454P.
XX PR 09-JUN-2000; 2000US-0211314P.
XX PR 18-JUL-2000; 2000US-0219007P.
XX PR 13-DEC-2000; 2000US-0255281P.
XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX PI Schlegel R, Endege WO, Monahan JE;
XX PF 2001-662795/76.
XX DR WPI; 2001-662795/76.
XX PT Novel isolated nucleic acid molecule associated with cancerous state of
XX PT prostate cells and correlating with presence of prostate cancer, useful
XX PT for detecting presence of prostate cancer, stage of prostate cancer.
XX PS Claim 1; Page 8784; 11750pp; English.
XX CC The invention relates to an isolated nucleic acid molecule (I) comprising
XX CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
XX CC specification or its complement. (I) is useful for: (a) assessing whether
XX CC a patient is afflicted with prostate cancer; (b) monitoring the
XX CC progression of prostate cancer in a patient; (c) assessing the efficacy
XX CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
XX CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
XX CC (e) selecting a composition for inhibiting prostate cancer in a patient;
XX CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
XX CC determining whether prostate cancer has metastasized in a patient; (h)
XX CC assessing the aggressiveness or indolence of prostate cancer in a patient
XX CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX SQ Sequence 459 BP; 120 A; 122 C; 111 G; 106 T; 0 U; 0 Other;
SQ Sequence 459 BP; 120 A; 122 C; 111 G; 106 T; 0 U; 0 Other;
Query Match 66.4%; Score 16.6; DB 5; Length 459;

Job time : 211.111 secs

Best Local Similarity 82.6%; Pred. No. 8.4e+02;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAACGACACAGCGCCCTC 23
||||| ||||| ||||| |||||
Db 82 AAAAAAGAGACACAGCGCTCCC 104

RESULT 50

ABV35393

ID ABV35393 standard; cDNA; 459 BP.

XX

AC ABV35393;

XX

DT 16-SEP-2002 (first entry)

XX

DE Human prostate expression marker cDNA 35384.

XX

KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;

XX

KW Pharmacogenomic marker; gene; ss.

XX

OS Homo sapiens.

XX

PN WO200160860-A2.

XX

PD 23-AUG-2001.

XX

PF 20-FEB-2001; 2001WO-US005171.

XX

PR 17-FEB-2000; 2000US-0183319P.

PR

PR 16-MAR-2000; 2000US-0189862P.

PR

PR 25-MAY-2000; 2000US-0207454P.

PR

PR 09-JUN-2000; 2000US-0211314P.

PR

PR 18-JUL-2000; 2000US-0219007P.

PR

PR 13-DEC-2000; 2000US-0255281P.

XX

PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX

PI Schlegel R, Endege WO, Monahan JE;

XX

DR WPI; 2001-662795/76.

XX

PT Novel isolated nucleic acid molecule associated with cancerous state of

PT

PT prostate cells and correlating with presence of prostate cancer, useful

PT

PT for detecting presence of prostate cancer, stage of prostate cancer.

XX

PS Claim 1; Page 7374; 11750pp; English.

XX

CC The invention relates to an isolated nucleic acid molecule (I) comprising

CC

CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the

CC

CC specification or its complement. (i) is useful for: (a) assessing whether

CC

CC a patient is afflicted with prostate cancer; (b) monitoring the

CC

CC progression of prostate cancer in a patient; (c) assessing the efficacy

CC

CC of a test compound to inhibit prostate cancer in a patient; (d) assessing

CC

CC the efficacy of a therapy for inhibiting prostate cancer in a patient;

CC

CC (e) selecting a composition for inhibiting prostate cancer in a patient;

CC

CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)

CC

CC determining whether prostate cancer has metastasized in a patient; (h)

CC

CC assessing the aggressiveness or indolence of prostate cancer in a patient

CC

CC; (I) is also useful as a pharmacodynamic or pharmacogenomic marker

SQ Sequence 459 BP; 120 A; 122 C; 111 G; 106 T; 0 U; 0 Other;

Query Match 66.4%; Score 16.6; DB 5; Length 459;

Best Local Similarity 82.6%; Pred. No. 8.4e+02;

Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAACGACACAGCGCCCTC 23
||||| ||||| ||||| |||||

Db 82 AAAAAAGAGACACAGCGCTCCC 104
||||| ||||| ||||| |||||

, Search completed: February 3, 2006, 21:56:49

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 3, 2006, 14:38:25 ; Search time 2951.67 Seconds
(without alignments)
396.277 Million cell updates/sec

Title: US-10-719-900-15
Perfect score: 25
Sequence: 1 aaaaacgacacagcagccctcgg 25

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues
Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 150 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_hic:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est7:*
9: gb_gss1:*
10: gb_gss2:*
11: gb_gss3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	23.4	93.6	229	2	BB288882
2	23.4	93.6	297	6	CA895258
3	23.4	93.6	310	1	AW411607
4	23.4	93.6	320	2	BE630623
5	23.4	93.6	322	1	AA790812
6	23.4	93.6	326	8	W36874
7	23.4	93.6	328	1	AA871673
8	23.4	93.6	330	1	AW647013
9	23.4	93.6	336	5	BY497898
10	23.4	93.6	336	2	BE959526
11	23.4	93.6	337	6	CB570825
12	23.4	93.6	341	1	AA066781
13	23.4	93.6	342	1	AI843447
14	23.4	93.6	346	1	AA080458
15	23.4	93.6	346	2	BF319961
16	23.4	93.6	348	2	BE629477
17	23.4	93.6	352	2	BB685033
18	23.4	93.6	353	5	BY398944
19	23.4	93.6	354	5	BY497353
20	23.4	93.6	355	2	BE197629
21	23.4	93.6	357	3	EM899372
22	23.4	93.6	358	1	AW412420

C	23	23.4	93.6	361	2	BE994381	UI-M-B21-	BE994381	UI-M-B21-
24	23.4	93.6	361	5	BY608470	BY608470	BY608470	BY608470	BY608470
C	25	23.4	93.6	366	2	BE989916	UI-M-B21-	BE989916	UI-M-B21-
26	23.4	93.6	369	1	AA396989	AA396989	AA396989	AA396989	AA396989
27	23.4	93.6	370	3	BI655761	BI655761	BI655761	BI655761	BI655761
28	23.4	93.6	370	5	BY504438	BY504438	BY504438	BY504438	BY504438
C	29	23.4	93.6	372	5	BF227139	UI-M-B21-	BF227139	UI-M-B21-
30	23.4	93.6	372	5	BY699023	BY699023	BY699023	BY699023	BY699023
31	23.4	93.6	374	1	AA396226	AA396226	AA396226	AA396226	AA396226
32	23.4	93.6	375	2	BB744103	BB744103	BB744103	BB744103	BB744103
33	23.4	93.6	378	2	BB736988	BB736988	BB736988	BB736988	BB736988
C	34	23.4	93.6	379	2	BF459322	UI-M-B21-	BF459322	UI-M-B21-
35	23.4	93.6	380	1	AA250176	AA250176	AA250176	AA250176	AA250176
C	36	23.4	93.6	381	1	AA088987	UI-M-B21-	AA088987	UI-M-B21-
37	23.4	93.6	381	1	AA240659	AA240659	AA240659	AA240659	AA240659
38	23.4	93.6	382	5	BY629510	BY629510	BY629510	BY629510	BY629510
39	23.4	93.6	382	5	BY685406	BY685406	BY685406	BY685406	BY685406
40	23.4	93.6	382	5	BF235341	BF235341	BF235341	BF235341	BF235341
41	23.4	93.6	383	2	BE197726	BE197726	BE197726	BE197726	BE197726
C	42	23.4	93.6	384	2	BE929024	UI-M-B21-	BE929024	UI-M-B21-
43	23.4	93.6	384	5	BY699498	BY699498	BY699498	BY699498	BY699498
44	23.4	93.6	384	5	BF320359	BF320359	BF320359	BF320359	BF320359
C	45	23.4	93.6	386	2	BY642036	UI-M-B21-	BY642036	UI-M-B21-
46	23.4	93.6	386	5	BY603881	BY603881	BY603881	BY603881	BY603881
47	23.4	93.6	387	5	BY630813	BY630813	BY630813	BY630813	BY630813
48	23.4	93.6	387	5	BY660055	BY660055	BY660055	BY660055	BY660055
49	23.4	93.6	388	2	BB737836	BB737836	BB737836	BB737836	BB737836
50	23.4	93.6	388	5	BY401120	BY401120	BY401120	BY401120	BY401120
C	51	23.4	93.6	389	5	BE994338	UI-M-B21-	BE994338	UI-M-B21-
52	23.4	93.6	392	2	BE994338	BE994338	BE994338	BE994338	BE994338
53	23.4	93.6	393	1	AA591164	AA591164	AA591164	AA591164	AA591164
54	23.4	93.6	393	5	BY427616	BY427616	BY427616	BY427616	BY427616
55	23.4	93.6	394	5	BY629324	BY629324	BY629324	BY629324	BY629324
56	23.4	93.6	395	5	BY438763	BY438763	BY438763	BY438763	BY438763
57	23.4	93.6	397	5	BY625494	BY625494	BY625494	BY625494	BY625494
58	23.4	93.6	397	5	BY627646	BY627646	BY627646	BY627646	BY627646
59	23.4	93.6	397	5	BY631643	BY631643	BY631643	BY631643	BY631643
60	23.4	93.6	398	5	BY575337	BY575337	BY575337	BY575337	BY575337
C	61	23.4	93.6	399	2	BE953304	UI-M-C31-	BE953304	UI-M-C31-
62	23.4	93.6	399	5	BY587583	BY587583	BY587583	BY587583	BY587583
C	63	23.4	93.6	399	6	CA893542	UI-M-B21-	CA893542	UI-M-B21-
64	23.4	93.6	400	3	BI794673	BI794673	BI794673	BI794673	BI794673
65	23.4	93.6	400	5	BY579116	BY579116	BY579116	BY579116	BY579116
66	23.4	93.6	401	2	BB742737	BB742737	BB742737	BB742737	BB742737
C	67	23.4	93.6	401	7	CK330945	H8249F02-	CK330945	H8249F02-
68	23.4	93.6	402	3	BQ265864	BQ265864	BQ265864	BQ265864	BQ265864
C	69	23.4	93.6	403	1	AW240772	UI-M-B21-	AW240772	UI-M-B21-
70	23.4	93.6	404	2	BE307297	BE307297	BE307297	BE307297	BE307297
71	23.4	93.6	404	5	BY601389	BY601389	BY601389	BY601389	BY601389
72	23.4	93.6	406	5	BY392307	BY392307	BY392307	BY392307	BY392307
73	23.4	93.6	406	5	BY638342	BY638342	BY638342	BY638342	BY638342
C	74	23.4	93.6	407	2	BE994642	UI-M-B21-	BE994642	UI-M-B21-
75	23.4	93.6	408	2	BB710131	BB710131	BB710131	BB710131	BB710131
76	23.4	93.6	408	5	BY511458	BY511458	BY511458	BY511458	BY511458
77	23.4	93.6	410	3	BI715778	BI715778	BI715778	BI715778	BI715778
C	78	23.4	93.6	411	2	BF321381	UI-M-B21-	BF321381	UI-M-B21-
79	23.4	93.6	411	5	BY439640	BY439640	BY439640	BY439640	BY439640
80	23.4	93.6	411	5	BY574665	BY574665	BY574665	BY574665	BY574665
81	23.4	93.6	411	5	BY626670	BY626670	BY626670	BY626670	BY626670
C	82	23.4	93.6	413	2	BF457487	UI-M-B21-	BF457487	UI-M-B21-
83	23.4	93.6	413	2	BB743431	BB743431	BB743431	BB743431	BB743431
84	23.4	93.6	414	2	BG061833	BG061833	BG061833	BG061833	BG061833
C	85	23.4	93.6	414	2	BB677028	UI-M-B21-	BB677028	UI-M-B21-
86	23.4	93.6	414	2	BB686470	BB686470	BB686470	BB686470	BB686470
C	87	23.4	93.6	414	3	BM203581	UI-M-B21-	BM203581	UI-M-B21-
88	23.4	93.6	414	5	BY649734	BY649734	BY649734	BY649734	BY649734
89	23.4	93.6	415	1	AW502341	AW502341	AW502341	AW502341	AW502341
C	90	23.4	93.6	416	1	AW763873	UI-M-B21-	AW763873	UI-M-B21-
91	23.4	93.6	416	2	BB828941	BB828941	BB828941	BB828941	BB828941
92	23.4	93.6	417	1	AA079500	AA079500	AA079500	AA079500	AA079500
C	93	23.4	93.6	420	2	BB742412	UI-M-B21-	BB742412	UI-M-B21-
94	23.4	93.6	421	2	BB677378	BB677378	BB677378	BB677378	BB677378
95	23.4	93.6	421	5	BY601791	BY601791	BY601791	BY601791	BY601791

BE994381	UI-M-BZ1-
BY608470	BY608470
BE989916	UI-M-BZ1-
AA396989	mx86c05.f
BI655761	603281362
BY504438	BY504438
BF227139	uz20505.x
BY699023	BY699023
AA396226	VB44H01.f
BB744103	BB744103
BB736988	BB736988
BF459322	UI-M-BZ1-
AA250176	mx85b12.f
AA088987	mm51905.f
AA240659	mv22502.f
BY629510	BY629510
BY685406	BY685406
BF235341	602036734
BE197726	u877c09.x
BY699498	BY699498
BF320359	u252h08.x
BY642036	BY642036
BY603881	BY603881
BY630813	BY630813
BY660055	BY660055
BB737836	BB737836
BY401120	BY401120
BE994338	UI-M-BZ1-
AA591164	vn53908.f
BY427616	BY427616
BY629324	BY629324
BY438763	BY438763
BY625494	BY625494
BY627646	BY627646
BY631643	BY631643
BY575337	BY575337
BE953304	UI-M-CD1-
BY587583	BY587583
CA893542	B0179003-
BI794673	1c63b07.x
BY579116	BY579116
BB742737	BB742737
CK330945	H8249F02-
BQ265864	NTSC_ff09
AW240772	uq33d08.x
BE307297	601091222
BY601389	BY601389
BY392307	BY392307
BY638342	BY638342
BE994642	UI-M-BZ1-
BB710131	BB710131
BY511458	BY511458
BI715778	1c63b07.y
BF321381	uz61f08.x
BY439640	BY439640
BY574665	BY574665
BY626670	BY626670
BF457487	UI-M-BZ1-
BB743431	BB743431
BG061833	L0951G04-
BB677028	BB677028
BB686470	BB686470
BM203581	C0256F03-
BY649734	BY649734
BY502341	BY502341
AW763873	uz66h09.x
BB828941	BB828941
AA079500	zm96h10.s
BB742412	BB742412
BB677378	BB677378
BY601791	BY601791


```

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Murinae; Mus.
1 (bases 1 to 320)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Other_ESTs: uu38d12.y1
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:1083835.

FEATURES
    source
        Location/Qualifiers
            1..320
                /organism="Mus musculus"
                /mol_type="mRNA"
                /db_xref="taxon:10090"
                /clone="IMAGE:3374231"
                /sex="female (lactating)"
                /tissue_type="mammary gland"
                /lab_host="DH10B"
                /clone_lib="Soares mammary gland NMLMG"
                /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                polylinker; 1st strand cDNA was prepared from mammary
                gland tissue from a lactating female, and was then primed
                with a Not I - oligo(dT) primer. Double-stranded cDNA was
                ligated to Eco RI adaptors (Pharmacia), digested with Not
                I and cloned into the Not I and Eco RI sites of the
                modified pT7T3 vector. Library is normalized. Library
                was constructed by Bento Soares and M. Fatima Bonaldo."
```

```

Query Match      93.6%; Score 23.4; DB 2; Length 320;
Best Local Similarity 96.0%; Pred. No. 5.8;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAAAACGACACGAGCGCCCTCGG 25
    |||||
Db 318 AAAAAACGACACTGCGAGCCCTCGG 294

RESULT 5
AA790812 322 bp mRNA linear EST 06-FEB-1998
LOCUS vW19d12.r1 Soares mammary_gland_NbMMG Mus musculus cDNA clone
DEFINITION IMAGE:1244279 5', mRNA sequence.
ACCESSION AA790812
VERSION AA790812.1 GI:2850932
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Murinae; Mus.
1 (bases 1 to 322)
Marras, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
Waterston, R.
The WashU-HHMI Mouse EST Project
Unpublished (1996)
Contact: Marra M/Mouse EST Project
WashU-HHMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:657967

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Murinae; Mus.
1 (bases 1 to 320)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Other_ESTs: uu38d12.y1
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:1083835.

FEATURES
    source
        Location/Qualifiers
            1..320
                /organism="Mus musculus"
                /mol_type="mRNA"
                /db_xref="taxon:10090"
                /clone="IMAGE:3374231"
                /sex="female (lactating)"
                /tissue_type="mammary gland"
                /lab_host="DH10B"
                /clone_lib="Soares mammary gland NMLMG"
                /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                polylinker; 1st strand cDNA was prepared from mammary
                gland tissue from a lactating female, and was then primed
                with a Not I - oligo(dT) primer. Double-stranded cDNA was
                ligated to Eco RI adaptors (Pharmacia), digested with Not
                I and cloned into the Not I and Eco RI sites of the
                modified pT7T3 vector. Library is normalized. Library
                was constructed by Bento Soares and M. Fatima Bonaldo."
```

```

Seq primer: -28ml3 rev2 ET from Amersham
High quality sequence stop: 311.
Location/Qualifiers
    1..322
        /organism="Mus musculus"
        /mol_type="mRNA"
        /strain="C57BL/6J"
        /db_xref="taxon:10090"
        /clone="IMAGE:1244279"
        /sex="male"
        /tissue_type="mammary gland"
        /dev_stage="4 weeks"
        /lab_host="DH10B"
        /clone_lib="Soares mammary_gland NbMMG"
        /note="Organ: mammary gland; Vector: pT7T3D-Pac
        (Pharmacia) with a modified polylinker; Site 1: Not I;
        Site 2: Eco RI; 1st strand cDNA was primed with a Not I -
        oligo(dT) primer [5',
        TGTTTACCAATCTGAAGTGGGCGCGCGAATGTTTTTTTTTTTTTTTTTTTT
        T 3']; double-stranded cDNA was ligated to Eco RI
        adaptors (Pharmacia), digested with Not I and cloned into
        the Not I and Eco RI sites of the modified pT7T3 vector.
        RNA provided by Dr. Minoru KO, Wayne State Univ. Library
        constructed and normalized by Bento Soares and M. Fatima
        Bonaldo."
```

```

ORIGIN
Query Match      93.6%; Score 23.4; DB 1; Length 322;
Best Local Similarity 96.0%; Pred. No. 5.8;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAAAACGACACGAGCGCCCTCGG 25
    |||||
Db 114 AAAAAACGACACTGCGAGCCCTCGG 138

RESULT 6
W36874 326 bp mRNA linear EST 11-SEP-1996
LOCUS mb77d10.r1 Soares mouse p3NMf19.5 Mus musculus cDNA clone
DEFINITION IMAGE:335443 5', mRNA sequence.
ACCESSION W36874
VERSION W36874.1 GI:1318149
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Murinae; Mus.
1 (bases 1 to 326)
Marras, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
Waterston, R.
The WashU-HHMI Mouse EST Project
Unpublished (1996)
Contact: Marra M/Mouse EST Project
WashU-HHMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:216843
Putative full length read
Seq primer: ETPrimer
High quality sequence stop: 321.
Location/Qualifiers
    1..326
        /organism="Mus musculus"
```

/mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="IMAGE:335443"
 /dev_stage="19.5 dpc total fetus"
 /lab_host="DH10B (ampicillin resistant)"
 /clone_lib="Soares mouse p3NM19.5"
 /note="Vector: pT73D (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', TGTTACGAATCTGAAGTCGGAGCGCGCCCTTTTTTTTTTTTTTTT 3']; double-stranded cDNA was ligated to Eco RI adaptors [AATTCGATCCCTG], digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Source irradiated bowel harvested 72 hours after irradiation (1400 Gys). Library constructed by Bob Barstead."

ORIGIN

Query Match 93.6%; Score 23.4; DB 8; Length 326;
 Best Local Similarity 96.0%; Pred. No. 5.8;
 Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAAAACGACACGCGAGCCCTCGG 25
 |||||
 DB 290 AAAAAACGACACTGCGAGCCCTCGG 314

RESULT 7

AA871673

LOCUS

DEFINITION

VQ39403.r1 Barstead bowel MPLRB9 Mus musculus cDNA clone

IMAGE:1096636 5', mRNA sequence.

AA871673

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished (1996)

Contact: Marra M/Mouse EST Project

WashU-HHMI Mouse EST Project

Washington University School of MedicineP

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: mouseest@wustl.wustl.edu

This clone is available royalty-free through LNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

MGI:602868

Putative full length read

vector to vector length is 517

Seq primer: -28ml3 rev2 ET from Amersham

High quality sequence stop: 297.

Location/Qualifiers

1..328

/organism="Mus musculus"

/mol_type="mRNA"

/strain="FVB/N"

/db_xref="taxon:10090"

/clone="IMAGE:1096636"

/tissue_type="bowel"

/dev_stage="8 weeks"

/lab_host="DH10B"

FEATURES

source

1..328

/organism="Mus musculus"

/mol_type="mRNA"

/strain="FVB/N"

/db_xref="taxon:10090"

/clone="IMAGE:1096636"

/tissue_type="bowel"

/dev_stage="8 weeks"

/lab_host="DH10B"

/clone_lib="Barstead bowel MPLRB9"
 /note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site 1: EcoRI; Site 2: NotI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', TGTTACGAATCTGAAGTCGGAGCGCGCCCTTTTTTTTTTTTTTTT 3']; double-stranded cDNA was ligated to Eco RI adaptors [AATTCGATCCCTG], digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Source irradiated bowel harvested 72 hours after irradiation (1400 Gys). Library constructed by Bob Barstead."

ORIGIN

Query Match 93.6%; Score 23.4; DB 1; Length 328;
 Best Local Similarity 96.0%; Pred. No. 5.8;
 Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAAAACGACACGCGAGCCCTCGG 25
 |||||
 DB 197 AAAAAACGACACTGCGAGCCCTCGG 221

RESULT 8

AW647013

LOCUS

DEFINITION

EST325603 R3DA Mus musculus cDNA clone R3DAB40, mRNA sequence.

AW647013

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished (1994)

Other ESTs: TC104313

Contact: Julie Earle-Hughes

TIGR

The Institute For Genomic Research

9712 Medical Center Drive, Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208

Location/Qualifiers

1..330

/organism="Mus musculus"

/mol_type="mRNA"

/cultivar="3T3"

/db_xref="taxon:10090"

/clone="R3DAB40"

/tissue_type="adipose"

/clone_lib="R3DA"

/note="Site_1: EcoRI; Site_2: XhoI"

FEATURES

source

1..330

/organism="Mus musculus"

/mol_type="mRNA"

/cultivar="3T3"

/db_xref="taxon:10090"

/clone="R3DAB40"

/tissue_type="adipose"

/clone_lib="R3DA"

/note="Site_1: EcoRI; Site_2: XhoI"

ORIGIN

Query Match 93.6%; Score 23.4; DB 1; Length 330;
 Best Local Similarity 96.0%; Pred. No. 5.8;
 Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAAAACGACACGCGAGCCCTCGG 25
 |||||
 DB 248 AAAAAACGACACTGCGAGCCCTCGG 272

RESULT 9

BY497898

LOCUS

DEFINITION

BY497898 RIKEN full-length enriched, bone marrow macrophage Mus

musculus cDNA clone 183052D10 3', mRNA sequence.

EST 14-DEC-2002

/dev stage="27-32 days"
 /lab_host="DH10B (Life Technologies)"
 /clone_lib="NIH_BMAP_MHI2_S1"
 /note="Vector: p773D-Pac (Pharmacia) with a modified
 polylinker; Site 1: Not 1; Site 2: Eco RI; The
 NIH_BMAP_MHI2_S1 library is a subtracted library derived
 from NIH_BMAP_MHI2. NIH_BMAP_MHI2 is a library derived
 from mouse hippocampus tissue. For a detailed description
 of the library from which this clone was derived, please
 visit our web site at brainest.eng.uiowa.edu.
 TAG_TISSUE=hippocampus
 TAG_LIB=NIH_BMAP_MHI2_S1
 TAG_SEQ=TAGCC"

ORIGIN

Query Match 93.6%; Score 23.4; DB 2; Length 336;
 Best Local Similarity 96.0%; Pred. No. 5.8;
 Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AAAAAACGACACGAGCCCTCGG 25
 |||||
 Db 306 AAAAAACGACACTCGAGCCCTCGG 282
 |||||

RESULT 11
 CBS70825
 LOCUS
 DEFINITION
 AGENCOURT 12976722 NIH_MGC_165 Mus musculus cDNA clone
 IMAGE:30278425 5', mRNA sequence.

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Mus musculus (house mouse)

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 NIH-MGC http://mgs.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgappe-remail.nih.gov

Tissue Procurement: Dr. Leslie L. Heckert
 cDNA Library Preparation: CLONTECH Laboratories, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: NDCM32 row: n column: 02
 High quality sequence stop: 337.
 High quality sequence stop: 337.

FEATURES

1. 337
 Location/Qualifiers
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="IMAGE:30278425"
 /tissue type="primary cultures of Sertoli cells"
 /lab_host="DH10B (TI-phage-resistant)"
 /clone_lib="NIH_MGC_165"
 /note="Organ: testis; Vector: pDNR-LIB; Site 1: SfiI
 (ggcattatggcc); Site 2: SfiI (ggcgcctcgcc); 5' and 3'
 adaptors were used in cloning as follows: 5' adaptor
 sequence: 5'-CACGCCATTATGGCC-3' and 3' adaptor sequence:
 5'-ATTCTAGAGCGCGAGCGCCGACATG-dt(30)BN-3' (where B = A,
 C, or G and N = A, C, G, or T). Average insert size 1.4 kb
 (range 0.6-3.5 kb). 15/15 colonies contained inserts by
 PCR. This library was enriched for full-length clones and
 was constructed by Clontech Laboratories (Palo Alto, CA).
 Note: this is a NIH_MGC Library."

ORIGIN

Query Match 93.6%; Score 23.4; DB 6; Length 337;
 Best Local Similarity 96.0%; Pred. No. 5.8;
 Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AAAAAACGACACGAGCCCTCGG 25
 |||||
 Db 22 AAAAAACGACACTCGAGCCCTCGG 46
 |||||

RESULT 12

AA066781
 LOCUS
 DEFINITION
 mm16e06.r1 Stratagene mouse diaphragm (#937303) Mus musculus cDNA
 clone IMAGE:521698 5', mRNA sequence.

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Mus musculus (house mouse)

REFERENCE
 AUTHORS
 Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
 Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
 Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
 Theising,B., Wyllie,T., Lennon,G., Soares,B., Wilson,R. and
 Waterston,R.
 The WashU-HMI Mouse EST Project
 Unpublished (1996)
 Contact: Marra M/Mouse EST Project
 WashU-HMI Mouse EST Project
 Washington University School of MedicineP
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu

TITLE
 JOURNAL
 COMMENT

IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:315546
 Seq primer: -28ml3 rev1 ET from Amersham
 High quality sequence stop: 300.
 High quality sequence stop: 300.

FEATURES

1. 341
 Location/Qualifiers
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="IMAGE:521698"
 /tissue type="diaphragm"
 /dev stage="adult"
 /lab_host="SOLR (kanamycin resistant)"
 /clone_lib="Stratagene mouse diaphragm (#937303)"
 /note="Organ: diaphragm; Vector: pBluescript SK-; Site 1:
 EcoRI; Site 2: XhoI; Cloned unidirectionally from mRNA
 prepared from diaphragm muscle. Primer: Oligo dt. Average
 insert size: 1.5 kb. Uni-ZAP XR Vector; -5' adaptor
 sequence: 5' GAATTCGGACGAG 3' -3' adaptor sequence: 5'
 CTCGACTTTTITTTTTTTT 3'."

ORIGIN

Query Match 93.6%; Score 23.4; DB 1; Length 341;
 Best Local Similarity 96.0%; Pred. No. 5.8;
 Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AAAAAACGACACGAGCCCTCGG 25
 |||||
 Db 274 AAAAAACGACACTCGAGCCCTCGG 298
 |||||

RESULT 13

AI843447
 LOCUS
 DEFINITION
 UI-M-AQ1-aec-d-11-0-UI.s1 NIH_BMAP_MHI_N Mus musculus cDNA clone
 EST 14-JUL-1999

UI-M-AQ1-aec-d-11-0-UI 3', mRNA sequence.

ACCESSION AI843447
 VERSION AI843447.1 GI:5477660
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 342)
 AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
 TITLE Normalization and subtraction: two approaches to facilitate gene discovery
 JOURNAL Genome Res. 6 (9), 791-806 (1996)
 PUBMED 8889548
 COMMENT Contact: Chin, H
 National Institute of Mental Health
 6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
 20892-9643, USA
 Tel: 301 443 1706
 Fax: 301 443 9890
 Email: mEST@mail.nih.gov

The sequence contained an oligo-dT track that was present in the oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A and the oligo-dT track served to verify it as a clone from the normalized hippocampus library cDNA library Preparation: M.B. Soares Lab Clone distribution: NIH BMAP cDNA clones will be made available by the means that is soon to be determined. When NIH determines the means for distributing of the BMAP cDNA clones, this record will be updated accordingly when that means is determined. The following repetitive elements were found in this cDNA sequence: 1-73, >AT rich#Low complexity
 Seq primer: M13 Forward
 POLYA=yes.

FEATURES
 source
 1..342
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UI-M-AQ1-aec-d-11-0-UI"
 /dev_stage="27-32 days"
 /lab_host="DH10B (Life Technologies)"
 /note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; The NIH_BMAP_MHI_n library is a normalized library constructed from mouse hippocampus. The tag is a string of 5 nucleotides present between the Not I site and the oligo-dT track. The library was constructed as described by Bonaldo, Lennon and Soares, Genome Research 6: 791-806, 1996. Tissue provided by Ms. Annie Novakovich, Zivic-Miller Laboratories.
 TAG_TISSUE=hippocampus
 TAG_LIB=NIH_BMAP_MHI_N
 TAG_SEQ=TTCCG"

ORIGIN
 Query Match 93.6%; Score 23.4; DB 1; Length 342;
 Best Local Similarity 96.0%; Pred. No. 5.8;
 Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 AAAAAACGACACGAGCGCCCTCGG 25
 |||||
 Db 326 AAAAAACGACACTGCGAGCCCTCGG 302
 |||||

RESULT 14
 AA080458 346 bp mRNA linear EST 15-FEB-1997
 LOCUS mn87g10.r1 Stratagene mouse Tcell 937311 Mus musculus cDNA clone
 DEFINITION

IMAGE:551106 5', mRNA sequence.

ACCESSION AA080458
 VERSION AA080458.1 GI:1619464
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 346)
 AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.
 TITLE The WashU-HHMI Mouse EST Project
 JOURNAL Unpublished (1996)
 COMMENT Contact: Marra M/Mouse EST Project
 WashU-HHMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu

This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:331898
 Seq primer: -28ml3 rev1 ET from Amer sham
 High quality sequence stop: 324.
 Location/Qualifiers
 1..346
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="IMAGE:551106"
 /tissue_type="Tcell"
 /dev_stage="M30 CD4+ cells"
 /lab_host="SOLR (kanamycin resistant)"
 /clone_lib="Stratagene mouse Tcell 937311"
 /note="Organ: blood; Vector: pBluescript SK-; Site_1: EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer: Oligo dt. M30 CD4+ cells. Average insert size: 1.0 kb; Uni-ZAP XR Vector; ~5' adaptor sequence: 5' GAATTCGACGAG 3' ~3' adaptor sequence: 5' CTCGAGTTTITTTTTTTT 3"

ORIGIN
 Query Match 93.6%; Score 23.4; DB 1; Length 346;
 Best Local Similarity 96.0%; Pred. No. 5.8;
 Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 AAAAAACGACACGAGCGCCCTCGG 25
 |||||
 Db 257 AAAAAACGACACTGCGAGCCCTCGG 233
 |||||

RESULT 15
 BF319961/c 346 bp mRNA linear EST 29-DEC-2000
 LOCUS uz44g01.x1 NCI_CGAP_Mam6 Mus musculus cDNA clone IMAGE:3671952 3', mRNA sequence.

ACCESSION BF319961
 VERSION BF319961.1 GI:11268813
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 346)
 AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
 JOURNAL Unpublished (1997)

COMMENT

Other ESTs: uz4g01.y1
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Jeffrey Green M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 image.llnl.gov/image/html/iresources.shtml

MGI:1432720

High quality sequence stop: 325.

FEATURES

Location/Qualifiers

source

1..346

/organism="Mus musculus"

/mol_type="mRNA"

/strain="FVB/N"

/db_xref="taxon:10090"

/clone="IMAGE:3671952"

/sex="female, virgin"

/tissue_type="infiltrating ductal carcinoma"

/dev_stage="5 months"

/lab_host="DH10B"

/clone_lib="NCI CGAP Mam6"

/note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;

Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.

Library constructed by Life Technologies. Investigator

providing samples: Jeffrey Green, M.D., NIH"

ORIGIN

Query Match 93.6%; Score 23.4; DB 2; Length 346;
 Best Local Similarity 96.0%; Pred. No. 5.8;
 Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAAAACGACACGAGCCCTCGG 25

|||||

Db 321 AAAAAACGACACTCGAGCCCTCGG 297

RESULT 16

BE629477

LOCUS

DEFINITION

un38d12.y1 Soares mammary_gland NMLMG Mus musculus cDNA clone

IMAGE:3374231 5', mRNA sequence.

BE629477.1 GI:9912165

EST.

KEYWORDS

SOURCE

ORGANISM

Mus musculus

Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

Sciurognathi; Muroidae; Murinae; Mus.

1 (bases 1 to 348)

NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

This clone is available royalty-free through LNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

MGI:1083835

Seq primer: -40RP from Gibco

High quality sequence stop: 311.

Location/Qualifiers

1..348

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

/clone="IMAGE:3374231"

/sex="female (lactating)"

/tissue_type="mammary gland"

/lab_host="DH10B"

/clone_lib="Soares mammary_gland NMLMG"

/note="Vector: pT7T3D-Pac (Pharmacia) with a modified

polylinker; 1st strand cDNA was prepared from mammary

gland tissue from a lactating female, and was then primed

with a Not I - oligo(dT) primer. Double-stranded cDNA was

ligated to Eco RI adaptors (Pharmacia), digested with Not

I and cloned into the Not I and Eco RI sites of the

modified pT7T3 vector. Library is normalized. Library

was constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 93.6%; Score 23.4; DB 2; Length 348;
 Best Local Similarity 96.0%; Pred. No. 5.8;
 Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAAAACGACACGAGCCCTCGG 25

|||||

Db 36 AAAAAACGACACTCGAGCCCTCGG 60

RESULT 17

BE685033

LOCUS

DEFINITION

mullerian duct Mus musculus cDNA clone 6820416K12 3', mRNA

sequence.

BE685033

VERSION

BE685033.1 GI:16011766

EST.

KEYWORDS

SOURCE

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

Sciurognathi; Muroidae; Murinae; Mus.

1 (bases 1 to 352)

Akimura,T., Arakawa,T., Carninci,P., Furuno,M., Hanagaki,T.,

Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Imotani,K.,

Ishii,Y., Ito,M., Kawai,J., Kojima,Y., Konno,H., Kouda,M.,

Katsuyama,T., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R.,

Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sakazume,N.,

Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,

Sogabe,Y., Suzuki,H., Tagawa,A., Takahashi,F., Takaku-Akai,H.,

Tanaka,T., Tomaru,A., Toya,T., Watahiki,A., Yasunishi,A.,

Muramatsu,M. and Hayashizaki,Y.

RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al.

2001)

Unpublished (2001)

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic

Sciences Center(GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Teurumi-Ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-res@sc.riken.jp, URL: http://genome.gsc.riken.jp/

Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,

Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new

genes. Genome Res. 10 (10), 1617-1630 (2000)

wagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,

Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T.,

Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A.

and Hayashizaki,Y.

RIKEN integrated sequence analysis (RISA) system--384-format

sequencing pipeline with 384 multicapillary sequencer. Genome Res.

10 (11), 1757-1771 (2000)

Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P.,

Sugahara,Y. and Hayashizaki,Y.

Computer-based methods for the mouse full-length cDNA

encyclopedia: real-time sequence clustering for construction of a

nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

VERSION BY497353.1 GI:26831732
 KEYWORDS Mus musculus (house mouse)
 SOURCE EST.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridea; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 354)
 AUTHORS Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S., Nikiado,I., Osato,N., Saito,K., Suzuki,H., Yamanaka,I., Kiyosawa,H., Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A., Schonbach,C., Gojobori,T., Baldarelli,R., Hill,D.P., Bult,C., Hume,D.A., Quackenbush,J., Schriml,L.M., Kanapin,A., Matsuda,H., Batalov,S., Beisel,K.W., Blake,J.A., Bradt,D., Brusic,V., Ciothia,C., Corbani,L.E., Cousins,S., Dalia,E., Dragani,T.A., Fletcher,C.F., Forrest,A., Frazer,K.S., Gaasterland,T., Gustincich,S., Hirokawa,N., Jackson,I.J., Jarvis,E.D., Kanai,A., Kawaji,H., Kawasawa,Y., Kedzierzki,R.M., King,B.L., Konagaya,A., Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R., Maltais,D., Marchionni,L., McKenzie,L., Miki,H., Nagashima,T., Numata,K., Okido,T., Pavan,W.J., Pertea,G., Pesole,G., Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramachandran,S., Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring,B.Z., Ringwald,M., Sandelin,A., Schneider,C., Semple,C.A., Setou,M., Shimada,K., Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tomita,M., Verardo,R., Wagner,L., Wahlstedt,C., Wang,Y., Watanabe,Y., Wells,C., Wilming,I.G., Wynshaw-Boris,A., Yanagisawa,M., Yang,I., Yang,L., Yuan,Z., Zavalon,M., Zhu,Y., Zimmer,A., Carninci,P., Hayateu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura,M., Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K., Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata,K., Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S., Rogers,J., Birney,E. and Hayashizaki,Y.
 TITLE Analysis of the mouse transcriptome based on functional annotation
 JOURNAL Nature 420, 563-573 (2002)
 PUBLISHED 12466851
 COMMENT Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@gs.riken.jp, URL:http://genome.gsc.riken.jp/
 Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hirozane,N., Imotani,K., Ishii,Y., Itoh,M., Kawai,J., Konno,H., Miyazaki,A., Murata,M., Nakamura,M., Nomura,K., Numazaki,R., Ohno,M., Sakai,K., Sakazume,N., Sasaki,D., Sato,K., Shibata,K., Shiraki,T., Tagami,M., Waki,K., Watahiki,A., Muramatsu,M. and Hayashizaki,Y. Direct Submission
 Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
 Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
 RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)
 Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
 Tissues were provided by David A. Hume (Depts. of Biochemistry and Microbiology/Parasitology Institute for Molecular Bioscience University of Queensland Brisbane, Q 4072 Australia) whose assistance we gratefully acknowledge.

Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.
 Location/Qualifiers
 1..354
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="I830049H02"
 /tissue_type="bone marrow"
 /cell_type="macrophage"
 /clone_lib="RIKEN full-length enriched, bone marrow macrophage"
 ORIGIN
 Query Match 93.6%; Score 23.4; DB 5; Length 354;
 Best Local Similarity 96.0%; Pred. No. 5.8;
 Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 AAAAAACGACACAGCGAGCCCTCGG 25
 |||||||||
 Db 40 AAAAAACGACACTGCGAGCCCTCGG 64
 |||||||||
 RESULT 20
 LOCUS BE197629/c
 DEFINITION BE197629
 ACCESSION BE197629
 VERSION BE197629.1 GI:8709798
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridea; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 355)
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
 Unpublished (1997)
 CONTACT: Robert Strausberg, Ph.D.
 Email: cgapbs@mail.nih.gov
 Tissue Procurement: Lothar Hennighausen Ph.D., Priscilla Furth Ph.D.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov/image/html/iresources.shtml
 MGI:1068426
 Possible reversed clone: polyT not found
 Seq primer: -40UP from Gibco
 High quality sequence stop: 347.
 Location/Qualifiers
 1..355
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="NMRI"
 /db_xref="taxon:10090"
 /clone="IMAGE:3257062"
 /tissue_type="tumor, gross tissue"
 /dev_stage="5 months"
 /lab_host="DH10B"
 /clone_lib="NCI CGAP Mam4"
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI; Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt. Library constructed by Life Technologies. Investigators providing samples: Lothar Hennighausen/Priscilla Furth, NIH Reference for transgenic model: Li et al., Cell Growth

and Differentiation 7, 3-11 (1996)."		TAG TISSUE=subfornical organ and postrema TAG_L1B=UI-M-DJ1 TAG_SEQ=GCTACATGAT"	
ORIGIN	Query Match Best Local Similarity 93.6%; Score 23.4; DB 2; Length 355; Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	Query Match Best Local Similarity 93.6%; Score 23.4; DB 3; Length 357; Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	
QY	1 AAAAAACGACACAGCGAGCCCTCGG 25 	1 AAAAAACGACACAGCGAGCCCTCGG 25 	
Db	288 AAAAAACGACACTGCGAGCCCTCGG 264	334 AAAAAACGACACTGCGAGCCCTCGG 310	
RESULT 21	BM899372 357 bp mRNA linear EST 12-MAR-2002	AW412420 358 bp mRNA linear EST 08-FEB-2000	
LOCUS	UI-M-DJ1-btq-c-05-0-UI.s1 NIH BMAP DJ1 Mus musculus cDNA clone	uq41b09.x1 NCI_CGAP_Mam5 Mus musculus cDNA clone IMAGE:2811929 3',	
DEFINITION	UI-M-DJ1-btq-c-05-0-UI 3', mRNA sequence.	mRNA sequence.	
ACCESSION	BM899372	AW412420	
VERSION	BM899372.1 GI:19382160	AW412420	
KEYWORDS	EST.	EST.	
SOURCE	Mus musculus (house mouse)	Mus musculus (house mouse)	
ORGANISM	Mus musculus	Mus musculus	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridea; Muridae; Murinae; Mus.	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridea; Muridae; Murinae; Mus.	
AUTHORS	Bonaldo,M.F., Lennon,G. and Soares,M.B.	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.	
TITLE	Normalization and subtraction: two approaches to facilitate gene discovery	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index	
JOURNAL	Genome Res. 6 (9), 791-806 (1996)	Unpublished (1997)	
PUBMED	8889548	Other_ESTs: uq41b09.y1	
COMMENT	Contact: Chin, H National Institute of Mental Health 6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD 20892-9643, USA Tel: 301 443 1706 Fax: 301 443 9890 Email: mEST@mail.nih.gov Tissue Procurement: Dr. Robin Davisson cDNA Library preparation: Dr. M. Bento Soares, University of Iowa cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa DNA Sequencing by: Dr. M. Bento Soares, University of Iowa Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com). The following repetitive elements were found in this cDNA sequence: 1-21, >AT rich#Low_complexity (matched complement) Seq primer: M13 FORWARD POLYA=Yes.	Contact: Robert Strausberg, Ph.D. Email: cgabbs-r@mail.nih.gov Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys cDNA Library preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Washington University Genome Sequencing Center Clone Distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html	
FEATURES	Location/Qualifiers	Location/Qualifiers	
source	1..357 /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6" /db_xref="taxon:10090" /clone="UI-M-DJ1-btq-c-05-0-UI" /tissue_type="subfornical organ and postrema" /dev_stage="Adult" /lab_host="DH10B (Life Technologies) (T1 phage resistant)" /clone_lib="NIH BMAP DJ1" /note="Organ: brain; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; NIH BMAP DJ1 is a cDNA library containing the following tissues: subfornical organ and postrema. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is GCTACATGAT."	1..358 /organism="Mus musculus" /mol_type="mRNA" /strain="mix FVB/N, C57BL/6J" /db_xref="taxon:10090" /clone="IMAGE:2811929" /tissue_type="tumor, gross tissue" /dev_stage="7 months" /lab_host="DH10B" /clone_lib="NCI CGAP Mam5" /note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI; Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT. Library constructed by Life Technologies. Investigators providing samples: Lothar Hennighausen/Robin Humphreys, NIH"	
ORIGIN	Query Match Best Local Similarity 93.6%; Score 23.4; DB 1; Length 358; Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	Query Match Best Local Similarity 96.0%; Pred. No. 5.8; Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	
QY	1 AAAAAACGACACAGCGAGCCCTCGG 25 	1 AAAAAACGACACAGCGAGCCCTCGG 25 	
Db	325 AAAAAACGACACTGCGAGCCCTCGG 301	325 AAAAAACGACACTGCGAGCCCTCGG 301	

RESULT 23	BE994381/c	361 bp	mRNA	linear	EST 05-OCT-2000
LOCUS	BE994381/c	361 bp	mRNA	linear	EST 05-OCT-2000
DEFINITION	UI-M-BZ1-bjm-e-10-0-UI.s1 NIH BMAP MH12 S1 Mus musculus cDNA clone	361 bp	mRNA	linear	EST 05-OCT-2000
ACCESSION	BE994381	361 bp	mRNA	linear	EST 05-OCT-2000
VERSION	BE994381.1	GI:10677563	mRNA	linear	EST 05-OCT-2000
KEYWORDS	EST.	GI:10677563	mRNA	linear	EST 05-OCT-2000
SOURCE	Mus musculus (house mouse)	GI:10677563	mRNA	linear	EST 05-OCT-2000
ORGANISM	Mus musculus	GI:10677563	mRNA	linear	EST 05-OCT-2000
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridea; Muridae; Murinae; Mus.	GI:10677563	mRNA	linear	EST 05-OCT-2000
AUTHORS	1 (bases 1 to 361)	GI:10677563	mRNA	linear	EST 05-OCT-2000
TITLE	Bonaldi,M.F., Lennon,G. and Soares,M.B.	GI:10677563	mRNA	linear	EST 05-OCT-2000
JOURNAL	Normalization and subtraction: two approaches to facilitate gene discovery	GI:10677563	mRNA	linear	EST 05-OCT-2000
PUBLISHED	Genome Res. 6 (9), 791-806 (1996)	GI:10677563	mRNA	linear	EST 05-OCT-2000
COMMENT	Contact: Chin, H National Institute of Mental Health 6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD 20892-9643, USA Tel: 301 443 1706 Fax: 301 443 9890 Email: mstet@mail.nih.gov The sequence contained an oligo-dT track that was present in the oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A tail. The sequence tag present in the cDNA between the NotI site and the oligo-dT track served to verify it as a clone from the hippocampus tissue cDNA Library Preparation: M.B. Soares Lab Clone distribution: Researchers may obtain BMAP cDNA clones from RESEARCH GENETICS. It should be noted that Bento Soares is generating a small number of additional specialized non-redundant arrays of BMAP cDNAs whose availability will be considered under appropriate and limited collaborative arrangements Seq primer: M13 Forward POLYA=Yes.	GI:10677563	mRNA	linear	EST 05-OCT-2000
FEATURES	Location/Qualifiers	GI:10677563	mRNA	linear	EST 05-OCT-2000
source	1. 361	GI:10677563	mRNA	linear	EST 05-OCT-2000
	/organism="Mus musculus"	GI:10677563	mRNA	linear	EST 05-OCT-2000
	/mol_type="mRNA"	GI:10677563	mRNA	linear	EST 05-OCT-2000
	/db_xref="taxon:10090"	GI:10677563	mRNA	linear	EST 05-OCT-2000
	/clone="UI-M-BZ1-bjm-e-10-0-UI"	GI:10677563	mRNA	linear	EST 05-OCT-2000
	/dev_stage="27-32 days"	GI:10677563	mRNA	linear	EST 05-OCT-2000
	/lab_host="DH10B (Life Technologies)"	GI:10677563	mRNA	linear	EST 05-OCT-2000
	/clone_lib="NIH BMAP MH12_S1"	GI:10677563	mRNA	linear	EST 05-OCT-2000
	/note="Vector: pRT3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; The NIH BMAP MH12_S1 library is a subtracted library derived from NIH BMAP MH12. NIH BMAP MH12 is a library derived from mouse hippocampus tissue. For a detailed description of the library from which this clone was derived, please visit our web site at brainest.eng.uiowa.edu."	GI:10677563	mRNA	linear	EST 05-OCT-2000
	TAG TISSUE=hippocampus	GI:10677563	mRNA	linear	EST 05-OCT-2000
	TAG LIB=NIH BMAP MH12_S1	GI:10677563	mRNA	linear	EST 05-OCT-2000
	TAG_SEQ=TAGTC	GI:10677563	mRNA	linear	EST 05-OCT-2000
ORIGIN		GI:10677563	mRNA	linear	EST 05-OCT-2000
Query Match	93.6%; Score 23.4; DB 2; Length 361;	GI:10677563	mRNA	linear	EST 05-OCT-2000
Best Local Similarity	96.0%; Pred. NO. 5.8;	GI:10677563	mRNA	linear	EST 05-OCT-2000
Matches	24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	GI:10677563	mRNA	linear	EST 05-OCT-2000
Qy	1 AAAAAACGACGACGAGCCCTCGG 25	GI:10677563	mRNA	linear	EST 05-OCT-2000
		GI:10677563	mRNA	linear	EST 05-OCT-2000
Db	333 AAAAAACGACGACGAGCCCTCGG 309	GI:10677563	mRNA	linear	EST 05-OCT-2000
		GI:10677563	mRNA	linear	EST 05-OCT-2000
RESULT 24	BY608470	361 bp	mRNA	linear	EST 15-DEC-2002
LOCUS	BY608470	361 bp	mRNA	linear	EST 15-DEC-2002
DEFINITION	CDNA clone K230302L24 3', mRNA sequence.	361 bp	mRNA	linear	EST 15-DEC-2002
ACCESSION	BY608470.1	GI:26943652	mRNA	linear	EST 15-DEC-2002
VERSION	BY608470.1	GI:26943652	mRNA	linear	EST 15-DEC-2002
KEYWORDS	EST.	GI:26943652	mRNA	linear	EST 15-DEC-2002
SOURCE	Mus musculus (house mouse)	GI:26943652	mRNA	linear	EST 15-DEC-2002
ORGANISM	Mus musculus	GI:26943652	mRNA	linear	EST 15-DEC-2002
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridea; Muridae; Murinae; Mus.	GI:26943652	mRNA	linear	EST 15-DEC-2002
AUTHORS	1 (bases 1 to 361)	GI:26943652	mRNA	linear	EST 15-DEC-2002
TITLE	Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S., Nikaido,I., Osato,N., Saito,R., Suzuki,H., Yamataka,I., Kiyosawa,H., Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A., Schonbach,C., Gojobori,T., Baldarelli,R., Hill,D.P., Bult,C., Hume,D.A., Quackenbush,J., Schriml,L.M., Kanapin,A., Matsuda,H., Batailo,S., Beisel,K.W., Blake,J.A., Bradt,D., Brusic,V., Chothia,C., Corbani,L.E., Cousins,S., Dalla,E., Dragani,T.A., Fletcher,C.F., Forrest,A., Frazer,K.S., Gaasterland,T., Gariboldi,M., Gissi,C., Godzik,A., Gough,J., Grimmond,S., Gustincich,S., Hirokawa,N., Jackson,I.J., Jarvis,E.D., Kanai,A., Kawaji,H., Kawasawa,Y., Kedzierski,R.M., King,B.L., Konagaya,A., Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R., Maltais,L., Marchionni,L., McKenzie,L., Miki,H., Nagashima,T., Numata,K., Okido,T., Pavan,W.J., Pertea,G., Pesole,G., Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramachandran,S., Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring,B.Z., Ringwald,M., Sandelin,A., Schneider,C., Semple,C.A., Setou,M., Shimada,K., Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tomita,M., Verardo,R., Wagner,L., Wahlestedt,C., Wang,Y., Watanabe,Y., Wells,C., Wilming,L.G., Wynshaw-Boris,A., Yanagisawa,M., Yang,I., Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A., Carninci,P., Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura,M., Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K., Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata,K., Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S., Rogers,J., Birney,E. and Hayashizaki,Y.	GI:26943652	mRNA	linear	EST 15-DEC-2002
	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs	GI:26943652	mRNA	linear	EST 15-DEC-2002
JOURNAL	Nature 420, 563-573 (2002)	GI:26943652	mRNA	linear	EST 15-DEC-2002
PUBLISHED	12466851	GI:26943652	mRNA	linear	EST 15-DEC-2002
COMMENT	Contact: Yoshihide Hayashizaki Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute The Institute of Physical and Chemical Research (RIKEN) 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan Tel: 81-45-503-9222 Fax: 81-45-503-9226 Email: genome-res@sc.riken.jp, URL: http://genome.gsc.riken.jp/ Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hirozane,T., Imotani,K., Ishii,Y., Itoh,M., Kawai,J., Konno,H., Miyazaki,A., Murata,M., Nakamura,M., Nomura,K., Numazaki,R., Ohno,M., Sakai,K., Sakazume,N., Sasaki,D., Sato,K., Shibata,K., Shiraki,T., Tagami,M., Waki,K., Watahiki,A., Muramatsu,M. and Hayashizaki,Y. Direct Submission Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001) Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000) RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000) Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001) cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.	GI:26943652	mRNA	linear	EST 15-DEC-2002

Tissues were provided by Michela Fagiolini and Takao K. Hensch (Laboratory for Neuronal Circuit Development Brain Science Institute RIKEN 2-1 Hirosewa,Wako-shi,Saitama 351-0198 Japan) whose assistance we gratefully acknowledge.
Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES
source

Location/Qualifiers

1. .361
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="K230302L24"
/tissue_type="visual cortex"
/clone_lib="RIKEN full-length enriched, visual cortex"

ORIGIN

Query Match 93.6%; Score 23.4; DB 5; Length 361;
Best Local Similarity 96.0%; Pred. No. 5.8;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAAAACGACACGAGCCCTCGG 25
|||||
Db 44 AAAAAACGACACTGCGAGCCCTCGG 68
|||||

RESULT 25
BE989916/c

LOCUS

DEFINITION UI-M-BZ1-bft-h-23-0-UI.s1 NIH_BMAP_MHI2_S1 Mus musculus cDNA clone

ACCESSION BE989916

VERSION BE989916

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridea; Muridae; Murinae; Mus.
1 (bases 1 to 366)

REFERENCE

AUTHORS Ronaldo,M.F., Lennon,G. and Soares,M.B.

TITLE Normalization and subtraction: two approaches to facilitate gene

discovery

Genome Res. 6 (9), 791-806 (1996)

GenBank

COMMENT Contact: Chin, H

National Institute of Mental Health

6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD

20892-9643, USA

Tel: 301 443 1706

Fax: 301 443 9890

Email: mEST@mail.nih.gov

The sequence contained an oligo-dT track that was present in the oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A tail. The sequence tag present in the cDNA between the NotI site and the oligo-dT track served to verify it as a clone from the hippocampus tissue cDNA Library Preparation: M.B. Soares Lab Clone distribution: Researchers may obtain BMAP cDNA clones from RESEARCH GENETICS. It should be noted that Bento Soares is generating a small number of additional specialized non-redundant arrays of BMAP cDNAs whose availability will be considered under appropriate and limited collaborative arrangements
Seq primer: M13 Forward
POLYA=Yes

FEATURES
source

Location/Qualifiers

1. .366
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UI-M-BZ1-bft-h-23-0-UI"
/dev_stage="27-32 days"

/lab_host="DH10B (Life Technologies)"

/clone_lib="NIH_BMAP_MHI2_S1"

/note="Vector: pT73D-Pac (Pharmacia) with a modified

polylinker; Site_1: Not I; Site_2: Eco RI; The

NIH_BMAP_MHI2_S1 library is a subtracted library derived

from NIH_BMAP_MHI2. NIH_BMAP_MHI2 is a library derived

from mouse hippocampus tissue. For a detailed description

of the library from which this clone was derived, please

visit our web site at brainest.eng.uiowa.edu.

TAG_TISSUE=hippocampus

TAG_LIB=NIH_BMAP_MHI2_S1

TAG_SEQ=GAGTC"

Query Match 93.6%; Score 23.4; DB 2; Length 366;

Best Local Similarity 96.0%; Pred. No. 5.8;

Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAAAACGACACGAGCCCTCGG 25
|||||

Db 333 AAAAAACGACACTGCGAGCCCTCGG 309
|||||

RESULT 26

AA396989

LOCUS

DEFINITION

AA396989

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Mus musculus

EST.

AA396989.1 GI:2050013

AA396989

AA396989

AA396989

AA396989

AA396989

AA396989

AA396989

AA396989

AA396989

AA396989

AA396989

AA396989

AA396989

AA396989

AA396989

AA396989

AA396989

AA396989

AA396989

AA396989

AA396989

AA396989

AA396989

AA396989

AA396989

AA396989

AA396989

AA396989

AA396989

AA396989

AA396989

AA396989

AA396989

AA396989

AA396989

AA396989

ORIGIN	Bonaldo."	musculus cDNA clone I830085H24 3', mRNA sequence.
Query Match	93.6%; Score 23.4; DB 1; Length 369;	
Best Local Similarity	96.0%; Pred. No. 5.8;	
Matches 24; Conservative	0; Mismatches 1; Indels 0; Gaps 0;	
QY	1 AAAAAACGACACAGCGAGCCCTCGG 25	
DB	256 AAAAAACGACACAGCGAGCCCTCGG 280	
RESULT 27		
BI655761		
LOCUS	370 bp mRNA linear EST 12-SEP-2001	
DEFINITION	603281362F1 NCI_CGAP_Mam4 Mus musculus cDNA clone IMAGE:5325831 5', mRNA sequence.	
ACCESSION	BI655761	
VERSION	BI655761.1 GI:155669997	
KEYWORDS	EST.	
SOURCE	Mus musculus (house mouse)	
ORGANISM	Mus musculus	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidae; Muridae; Murinae; Mus.	
AUTHORS	1 (bases 1 to 370)	
TITLE	NIH-MGC http://mgc.nci.nih.gov/	
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)	
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-x@mail.nih.gov Tissue Procurement: Lothar Hennighausen Ph.D., Priscilla Furth Ph.D. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: http://image.lnl.gov Plate: L1AM11826 row: k column: 16 High quality sequence stop: 370.	
FEATURES	Location/Qualifiers	
source	1..370	
	/organism="Mus musculus"	
	/mol_type="mRNA"	
	/strain="NMRI"	
	/db_xref="taxon:10090"	
	/clone="IMAGE:5325831"	
	/tissue_type="tumor, gross tissue"	
	/dev_stage="5 months"	
	/lab_host="DH10B"	
	/clone_lib="NCI_CGAP Mam4"	
	/note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI; Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT. Library constructed by Life Technologies. Investigators providing samples: Lothar Hennighausen/Priscilla Furth, NIH Reference for transgenic model: Li et al., Cell Growth and Differentiation 7, 3-11 (1996)."	
ORIGIN		
Query Match	93.6%; Score 23.4; DB 3; Length 370;	
Best Local Similarity	96.0%; Pred. No. 5.8;	
Matches 24; Conservative	0; Mismatches 1; Indels 0; Gaps 0;	
QY	1 AAAAAACGACACAGCGAGCCCTCGG 25	
DB	45 AAAAAACGACACAGCGAGCCCTCGG 69	
RESULT 28		
BY504438		
LOCUS	370 bp mRNA linear EST 14-DEC-2002	
DEFINITION	RIKEN full-length enriched, bone marrow macrophage Mus	

ACCESSION	BY504438	
VERSION	BY504438.1 GI:26838817	
KEYWORDS	EST.	
SOURCE	Mus musculus (house mouse)	
ORGANISM	Mus musculus	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidae; Muridae; Murinae; Mus.	
AUTHORS	1 (bases 1 to 370)	
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs	
JOURNAL	Nature 420, 563-573 (2002)	
PUBLISHED	12466851	
COMMENT	Contact: Yoshihide Hayashizaki Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute The Institute of Physical and Chemical Research (RIKEN) 1-7-22 Suehiro-cho, Teurumi-ku, Yokohama, Kanagawa 230-0045, Japan Tel: 81-45-503-9222 Fax: 81-45-503-9216 Email: genome-res@gsr.riken.jp, URL: http://genome.gsc.riken.jp/ Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ono, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001) Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000) RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000) Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001) cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Tissues were provided by David A. Hume (Depts. of Biochemistry and Microbiology/Parasitology Institute for Molecular Bioscience	

University of Queensland Brisbane, Q 4072 Australia) whose assistance we gratefully acknowledge.
Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES

Location/Qualifiers
1..370
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CS7BL/6J"
/db_xref="taxon:10090"
/clone="I830085H24"
/tissue_type="bone marrow"
/cell_type="macrophage"
/clone_lib="RIKEN full-length enriched, bone marrow macrophage"

ORIGIN

Query Match 93.6%; Score 23.4; DB 5; Length 370;
Best Local Similarity 96.0%; Pred. No. 5.8;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAAAACGACACGAGCCCTCGG 25
|||||
45 AAAAAACGACACTGGCGCCCTCGG 69

RESULT 29

BF227139/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

BF227139 372 bp mRNA linear EST 29-DRC-2000
uz20b05.x1 NCI_CGAP_Mam5 Mus musculus cDNA clone IMAGE:3669585 3',
mRNA sequence.
BF227139
BF227139.1 GI:11135043
EST.
Mus musculus (house mouse)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Murinae; Mus.
1 (bases 1 to 372)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Other ESTs: uz20b05.y1
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov

JOURNAL

COMMENT

Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone Distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
image.llnl.gov/image/html/iresources.shtml

MG1:1430353

High quality sequence stop: 366.

FEATURES

source

Location/Qualifiers
1..372
/organism="Mus musculus"
/mol_type="mRNA"
/strain="mix FVB/N, CS7BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:3669585"
/tissue_type="tumor, gross tissue"
/dev_stage="7 months"
/lab_host="DH10B"
/clone_lib="NCI_CGAP Mam5"
/notes="Organ: mammary; Vector: pCMV-SPORT6; Site 1: Salt;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Library constructed by Life Technologies. Investigators
providing samples: Lothar Hennighausen/Robin Humphreys,
NIH"

ORIGIN

Query Match 93.6%; Score 23.4; DB 2; Length 372;
Best Local Similarity 96.0%; Pred. No. 5.8;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAAAACGACACGAGCCCTCGG 25
|||||
316 AAAAAACGACACTGGCGCCCTCGG 292

RESULT 30

BF227139

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

BF227139 372 bp mRNA linear EST 16-DEC-2002
BY699023 RIKEN full-length enriched, osteoclast-like cell Mus
musculus cDNA clone I420114J10 3', mRNA sequence.
BY699023
BY699023.1 GI:27110050
EST.
Mus musculus (house mouse)

REFERENCE

AUTHORS

1 (bases 1 to 372)
Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
Nikaido, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I.,
Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,
Schonbach, C., Gotohori, T., Baldarelli, R., Hill, D.P., Bult, C.,
Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H.,
Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusic, V.,
Chothia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A.,
Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T.,
Gariboldi, M., Gibsi, C., Godzik, A., Gough, J., Grimmond, S.,
Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A.,
Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A.,
Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R.,
Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T.,
Numata, K., Okido, T., Pavan, W.J., Perlea, G., Pesole, G.,
Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S.,
Ravasi, T., Reed, J.C., Reid, J., Ring, B.Z., Ringwald, M.,
Sandelin, A., Schneider, C., Semple, C.A., Setou, M., Shimada, K.,
Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M.,
Verardo, R., Wagner, L., Wahlstedt, C., Wang, Y., Watanabe, Y.,
Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I.,
Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P.,
Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M.,
Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,
Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y.,
Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K.,
Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S.,
Rogers, J., Birney, E. and Hayashizaki, Y.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)

JOURNAL

PUBMED

COMMENT

Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.jp, <http://genome.gsc.riken.jp/>
Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S.,
Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H.,
Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R.,
Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K.,
Shiraki, T., Tsgami, M., Waki, K., Watahiki, A., Muramatsu, M. and
Hayashizaki, Y. Direct Submission
Computational Analysis of Full-Length Mouse cDNAs Compared with
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Tissues were provided by Takashi Ishikawa (Department of Surgery 2 Yokohama City University 3-9 Fukuura, Kanazawa-ku, Yokohama 236-0004 Japan) whose assistance we gratefully acknowledge.
Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES

Location/Qualifiers
1. .372
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="I420114J10"
/cell_type="osteoclast-like cell"
/clone_lib="RIKEN full-length enriched, osteoclast-like cell"

ORIGIN

Query Match 93.6%; Score 23.4; DB 5; Length 372;
Best Local Similarity 96.0%; Pred. No. 5.8;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAAAACGACACGCGCGCCTCGG 25

|||||
59 AAAAAACGACACGCGCGCCTCGG 83

RESULT 31

AA396226

LOCUS

DEFINITION v044h01.r1 Soares mouse lymph node NbMLN Mus musculus cDNA clone IMAGE:751825 5', mRNA sequence.

ACCESSION

AA396226

VERSION

AA396226.1

KEYWORDS

EST.

SOURCE

Mus musculus

ORGANISM

Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

Sciurognathi; Muroidea; Muridae; Murinae; Mus.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Contact: Marra M/Mouse EST Project

WashU-HMI Mouse EST Project

Washington University School of MedicineP

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: mouseest@wustl.edu

This clone is available royalty-free through LLNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

MGI:460809

Seq primer: -28m13 rev2 ET from Amersham.

Location/Qualifiers

1. .374

.. source

/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:751825"
/sex="male"
/tissue_type="lymph node"
/dev_stage="4 weeks"
/lab_host="DH10B"
/clone_lib="Soares mouse lymph node NbMLN"
/notes="Organ: lymph node; Vector: pT7D-Pac (Pharmacia) with a modified polylinker; Site_1: Not 1; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5]."

TGTTCAATCTCAAGTCGAGCGCGGATCTTTTTTTTTTTTTTTTTTTTTTTT
3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia) digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. RNA provided by Dr. Bertrand Jordan. Library constructed and normalized by Bento Soares and M.Fatima Bonaldo."

ORIGIN

Query Match 93.6%; Score 23.4; DB 1; Length 374;
Best Local Similarity 96.0%; Pred. No. 5.8;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAAAACGACACGCGCGCCTCGG 25

|||||
111 AAAAAACGACACGCGCGCCTCGG 135

RESULT 32

BB744103

LOCUS

DEFINITION

BB744103 RIKEN full-length enriched, adult male kidney Mus musculus cDNA clone F520015N21 3', mRNA sequence.

ACCESSION

BB744103

VERSION

BB744103.1

KEYWORDS

EST.

SOURCE

Mus musculus

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

Sciurognathi; Muroidea; Muridae; Murinae; Mus.

REFERENCE

AUTHORS

1 (bases 1 to 375)

Akimura,T., Arakawa,T., Carninci,P., Furuno,M., Hanagaki,T., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Imotani,K., Ishii,Y., Ito,M., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Matsuyama,T., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sakazume,N., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Watanishi,A., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.

RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura.T., et al. 2001)

Unpublished (2001)

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic

Sciences Center (GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-23 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9226

Email: genome-res@sc.riken.jp, URL: <http://genome.gsc.riken.jp/>

Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new

genes. Genome Res. 10 (10), 1617-1630 (2000)

wagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watanishi,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A.

20892-9643, USA
Tel: 301 443 1706
Fax: 301 443 9890
Email: mEst@mail.nih.gov

The sequence contained an oligo-dT track that was present in the oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A tail. The sequence tag present in the cDNA between the NotI site and the oligo-dT track served to verify it as a clone from the hippocampus tissue cDNA library Preparation: M.B. Soares Lab Clone distribution: Researchers may obtain BMAP cDNA clones from RESEARCH GENETICS. It should be noted that Bento Soares is generating a small number of additional specialized non-redundant arrays of BMAP cDNAs whose availability will be considered under appropriate and limited collaborative arrangements. The following repetitive elements were found in this cDNA sequence: 1-21,
>AT-rich#Low complexity
Seq primer: M13 Forward
POLYA=Yes.

FEATURES

source Location/Qualifiers
1..379
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UI-M-BZ1-bmc-a-08-0-UI"
/dev_stage="27-32 days"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NIH_BMAP_MHI2_S1"
/notes="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; The NIH_BMAP_MHI2_S1 library is a subtracted library derived from NIH_BMAP_MHI2. NIH_BMAP_MHI2 is a library derived from mouse hippocampus tissue. For a detailed description of the library from which this clone was derived, please visit our web site at brainest.eng.uiowa.edu.
TAG_TISSUE=hippocampus
TAG_LIB=NIH_BMAP_MHI2_S1
TAG_SEQ=TAGTC"

ORIGIN

Query Match 93.6%; Score 23.4; DB 2; Length 379;
Best Local Similarity 96.0%; Pred. No. 5.8;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 AAAAAACGACACGAGCGCCCTCGG 25
|||||
Db 334 AAAAAACGACACTGCGAGCCCTCGG 310

RESULT 35

AA250176
LOCUS mx85b12.r1 Soares mouse NML Mus musculus cDNA clone IMAGE:693119
DEFINITION 5', mRNA sequence.
ACCESSION AA250176
VERSION AA250176.1 GI:1882614
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM

REFERENCE

AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.
TITLE The WashU-HMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project

20892-9643, USA
Tel: 301 443 1706
Fax: 301 443 9890
Email: mEst@mail.nih.gov

The sequence contained an oligo-dT track that was present in the oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A tail. The sequence tag present in the cDNA between the NotI site and the oligo-dT track served to verify it as a clone from the hippocampus tissue cDNA library Preparation: M.B. Soares Lab Clone distribution: Researchers may obtain BMAP cDNA clones from RESEARCH GENETICS. It should be noted that Bento Soares is generating a small number of additional specialized non-redundant arrays of BMAP cDNAs whose availability will be considered under appropriate and limited collaborative arrangements. The following repetitive elements were found in this cDNA sequence: 1-21,
>AT-rich#Low complexity
Seq primer: M13 Forward
POLYA=Yes.

FEATURES

source Location/Qualifiers
1..379
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UI-M-BZ1-bmc-a-08-0-UI"
/dev_stage="27-32 days"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NIH_BMAP_MHI2_S1"
/notes="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; The NIH_BMAP_MHI2_S1 library is a subtracted library derived from NIH_BMAP_MHI2. NIH_BMAP_MHI2 is a library derived from mouse hippocampus tissue. For a detailed description of the library from which this clone was derived, please visit our web site at brainest.eng.uiowa.edu.
TAG_TISSUE=hippocampus
TAG_LIB=NIH_BMAP_MHI2_S1
TAG_SEQ=TAGTC"

ORIGIN

Query Match 93.6%; Score 23.4; DB 2; Length 379;
Best Local Similarity 96.0%; Pred. No. 5.8;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 AAAAAACGACACGAGCGCCCTCGG 25
|||||
Db 334 AAAAAACGACACTGCGAGCCCTCGG 310

RESULT 35

AA250176
LOCUS mx85b12.r1 Soares mouse NML Mus musculus cDNA clone IMAGE:693119
DEFINITION 5', mRNA sequence.
ACCESSION AA250176
VERSION AA250176.1 GI:1882614
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM

REFERENCE

AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.
TITLE The WashU-HMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project

Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu

This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:426679

Seq primer: -28ml3 rev2 ET from Amersham
High quality sequence stop: 379.

FEATURES

source Location/Qualifiers
1..380
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:693119"
/tissue_type="Liver"
/lab_host="DH10B"
/clone_lib="Soares mouse NML"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', TGTTCACCAATCTGAAGTGGACGGCGCGAATCTTTTTCCTTTT 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library constructed and normalized by Bento Soares and M.Fatima Bonaldo."

ORIGIN

Query Match 93.6%; Score 23.4; DB 1; Length 380;
Best Local Similarity 96.0%; Pred. No. 5.8;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 AAAAAACGACACGAGCGCCCTCGG 25
|||||
Db 268 AAAAAACGACACTGCGAGCCCTCGG 292

RESULT 36

AA088987/c
LOCUS mm51g05.r1 Stratagene mouse melanoma (#937312) Mus musculus cDNA
DEFINITION clone IMAGE:525080 5', mRNA sequence.
ACCESSION AA088987
VERSION AA088987.1 GI:1634457
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM

REFERENCE

AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.
TITLE The WashU-HMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu

TITLE

The WashU-HMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:318928
Seq primer: -28ml3 rev1 ET from Amersham
High quality sequence stop: 325.
Location/Qualifiers

```
source
1. 381
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:525080"
/tissue_type="melanoma"
/dev_stage="M2 cells"
/lab_host="SOLR (kanamycin resistant)"
/clone_lib="Stratagene mouse melanoma (#937312)"
EcoRI; Site 2: XhoI; Vector: pBluescript SK-; Site 1:
Oligo dt. Fom M2 cells, a highly metaetatic derivative of
the K-1735 (mouse) melanoma. Average insert size: 1.0 kb;
Uni-ZAP XR Vector; ~5' adaptor sequence: 5' GAATTCGGCAGG
3' ~3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTT 3"

ORIGIN
Query Match          93.6%; Score 23.4; DB 1; Length 381;
Best Local Similarity 96.0%; Pred. No. 5.8;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAAAACGACACGAGCCCTCGG 25
|||||
Db 234 AAAAAACGACACTGCGAGCCCTCGG 258
|||||

RESULT 38
BB737665
LOCUS
DEFINITION
Mus musculus
ACCESSION
BB737665
VERSION
BB737665.1
KEYWORDS
GI:16136815
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus

REFERENCE
AUTHORS
Akimura,T., Arakawa,T., Carninci,P., Furuno,M., Hanagaki,T.,
Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Imotani,K.,
Ishii,Y., Ito,M., Kawai,J., Kojima,Y., Konno,H., Kouda,M.,
Matsuyama,T., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R.,
Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sakazume,N.,
Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Suzuki,H., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
Tanaka,T., Tomaru,A., Toya,T., Watahiki,A., Yasunishi,A.,
Muramatsu,M. and Hayashizaki,Y.
RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al.
2001)
Unpublished (2001)
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
wagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,B.,
Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T.,
Matsura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A.
and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P.,
Sugahara,Y. and Hayashizaki,Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Please visit our web site (http://genome.gsc.riken.go.jp) for
further details.
e mouse tissues.
Location/Qualifiers
1. .382
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"

source
1. 381
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:525080"
/tissue_type="melanoma"
/dev_stage="M2 cells"
/lab_host="SOLR (kanamycin resistant)"
/clone_lib="Stratagene mouse melanoma (#937312)"
EcoRI; Site 2: XhoI; Vector: pBluescript SK-; Site 1:
Oligo dt. Fom M2 cells, a highly metaetatic derivative of
the K-1735 (mouse) melanoma. Average insert size: 1.0 kb;
Uni-ZAP XR Vector; ~5' adaptor sequence: 5' GAATTCGGCAGG
3' ~3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTT 3"

ORIGIN
Query Match          93.6%; Score 23.4; DB 1; Length 381;
Best Local Similarity 96.0%; Pred. No. 5.8;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAAAACGACACGAGCCCTCGG 25
|||||
Db 318 AAAAAACGACACTGCGAGCCCTCGG 294
|||||

RESULT 37
AA240659
LOCUS
DEFINITION
mv22c02.r1 GuayWoodford Beier mouse kidney day 0 Mus musculus cDNA
Clone IMAGE:655778 5', mRNA sequence.

ACCESSION
AA240659
VERSION
AA240659.1
KEYWORDS
GI:1865033
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus

REFERENCE
AUTHORS
Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
The WashU-HMI Mouse EST Project
Unpublished (1996)
Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LML; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:401626
Seq primer: -28ml3 rev1 ET from Amersham
High quality sequence stop: 351.
Location/Qualifiers
1. 381
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:655778"
/tissue_type="kidney"
/dev_stage="newborn (day 0)"
/lab_host="SOLR (kanamycin resistant)"
/clone_lib="GuayWoodford Beier mouse kidney day 0"
EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer:
Oligo dt. Average insert size: 1.0 kb; Uni-ZAP XR Vector;
```

```

/db xref="taxon:10090"
/clone="F430014L12"
/tissue_type="spleen"
/dev_stage="6 days neonate"
/clone_lib="RIKEN full-length enriched, 6 days neonate
spleen"

ORIGIN

Query Match          93.6%; Score 23.4; DB 2; Length 382;
Best Local Similarity 96.0%; Pred. No. 5.8;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAAAACGACACGCGCGCCTCGG 25
    |||||
Db 67 AAAAAACGACACTGCGCGCCTCGG 91

RESULT 39
BY629510
LOCUS
DEFINITION
cDNA clone K430031D24 3', mRNA sequence.
ACCESSION
BY629510
VERSION
EST.
KEYWORDS
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Murinae; Mus.
1 (bases 1 to 382)
Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S.,
Nikaido,I., Osato,N., Saito,R., Suzuki,H., Yamanaka,I.,
Kiyosawa,H., Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A.,
Schonbach,C., Gojobori,T., Baldarelli,R., Hill,D.P., Bult,C.,
Hume,D.A., Quackenbush,J., Schriml,L.M., Kanapin,A., Matsuda,H.,
Batalov,S., Beisel,K.W., Blake,J.A., Bradt,D., Brusic,V.,
Chothia,C., Corbani,L.E., Cousins,S., Dalia,E., Dragani,T.A.,
Chothia,C., Corbani,L.E., Cousins,S., Dalia,E., Dragani,T.A.,
Fletcher,C.F., Forrest,A., Frazer,K.S., Gaasterland,T.,
Gustincich,S., Godzik,A., Gough,J., Grimmond,S.,
Gariboldi,M., Gissi,C., Godzik,A., Gough,J., Grimmond,S.,
Gustincich,S., Godzik,A., Gough,J., Grimmond,S.,
Gustincich,S., Godzik,A., Gough,J., Grimmond,S.,
Kawaji,H., Kawanawa,Y., Kedzierski,R.M., King,B.L., Konagaya,A.,
Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R.,
Maltais,L., Marchionni,L., McKenzie,L., Miki,H., Nagashima,T.,
Numata,K., Okido,T., Pavan,W.J., Pertea,G., Pesole,G.,
Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramachandran,S.,
Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring,B.Z., Ringwald,M.,
Santelin,A., Schneider,C., Semple,C.A., Setou,M., Shimada,K.,
Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tomita,M.,
Verardo,R., Wagner,L., Wahlstedt,C., Wang,Y., Watanabe,Y.,
Wells,C., Wilming,L.G., Wynshaw-Boris,A., Yanagisawa,M., Yang,I.,
Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A., Carninci,P.,
Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura,M.,
Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K.,
Arakawa,T., Fukuda,S., Hara,A., Haehizume,W., Imotani,K., Ishii,Y.,
Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata,K.,
Shingawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S.,
Rogers,J., Birney,B. and Hayashizaki,Y.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
12466851
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/
Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S.,
Hirozane,T., Imotani,K., Ishii,Y., Itoh,M., Kawai,J., Konno,H.,
Miyazaki,A., Murata,M., Nakamura,M., Nomura,K., Numazaki,R.,

```

```

Ohno,M., Sakai,K., Sakazume,N., Sasaki,D., Sato,K., Shibata,K.,
Shiraki,T., Tagami,M., Waki,K., Watahiki,A., Muramatsu,M. and
Hayashizaki,Y. Direct Submission
Computational Analysis of Full-Length Mouse cDNAs Compared with
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN,
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Tissues were provided by Michela Fagiolini and Takao K. Hensch (
Laboratory for Neuronal Circuit Development Brain Science Institute
RIKEN 2-1 Hirosawa,Wako-shi,Saitama 351-0198 Japan ) whose
assistance we gratefully acknowledge.
Please visit our web site (http://genome.gsc.riken.go.jp) for
further details.

FEATURES             Location/Qualifiers
     1..382
        /organism="Mus musculus"
        /mol_type="mRNA"
        /strain="C57BL/6J"
        /db_xref="taxon:10090"
        /clone="K430031D24"
        /tissue_type="visual cortex"
        /clone_lib="RIKEN full-length enriched, visual cortex"

ORIGIN

Query Match          93.6%; Score 23.4; DB 5; Length 382;
Best Local Similarity 96.0%; Pred. No. 5.8;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAAAACGACACGCGCGCCTCGG 25
    |||||
Db 62 AAAAAACGACACTGCGCGCCTCGG 86

RESULT 40
BY685406
LOCUS
DEFINITION
cDNA clone K430011M17 3', mRNA sequence.
ACCESSION
BY685406
VERSION
EST.
KEYWORDS
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Murinae; Mus.
1 (bases 1 to 382)
Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S.,
Nikaido,I., Osato,N., Saito,R., Suzuki,H., Yamanaka,I.,
Kiyosawa,H., Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A.,
Schonbach,C., Gojobori,T., Baldarelli,R., Hill,D.P., Bult,C.,
Hume,D.A., Quackenbush,J., Schriml,L.M., Kanapin,A., Matsuda,H.,
Batalov,S., Beisel,K.W., Blake,J.A., Bradt,D., Brusic,V.,
Chothia,C., Corbani,L.E., Cousins,S., Dalia,E., Dragani,T.A.,
Chothia,C., Corbani,L.E., Cousins,S., Dalia,E., Dragani,T.A.,
Fletcher,C.F., Forrest,A., Frazer,K.S., Gaasterland,T.,
Gariboldi,M., Gissi,C., Godzik,A., Gough,J., Grimmond,S.,
Gustincich,S., Godzik,A., Gough,J., Grimmond,S.,
Gustincich,S., Godzik,A., Gough,J., Grimmond,S.,
Kawaji,H., Kawanawa,Y., Kedzierski,R.M., King,B.L., Konagaya,A.,
Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R.,
Maltais,L., Marchionni,L., McKenzie,L., Miki,H., Nagashima,T.,
Numata,K., Okido,T., Pavan,W.J., Pertea,G., Pesole,G.,
Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramachandran,S.,
Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring,B.Z., Ringwald,M.,
Santelin,A., Schneider,C., Semple,C.A., Setou,M., Shimada,K.,
Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tomita,M.,
Verardo,R., Wagner,L., Wahlstedt,C., Wang,Y., Watanabe,Y.,
Wells,C., Wilming,L.G., Wynshaw-Boris,A., Yanagisawa,M., Yang,I.,
Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A., Carninci,P.,
Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura,M.,
Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K.,
Arakawa,T., Fukuda,S., Hara,A., Haehizume,W., Imotani,K., Ishii,Y.,
Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata,K.,
Shingawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S.,
Rogers,J., Birney,B. and Hayashizaki,Y.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
12466851
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/
Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S.,
Hirozane,T., Imotani,K., Ishii,Y., Itoh,M., Kawai,J., Konno,H.,
Miyazaki,A., Murata,M., Nakamura,M., Nomura,K., Numazaki,R.,

```

Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramachandran,S., Ravasi,T., Reed,J.C., Reid,J., Reid,J., Ring,B.Z., Ringwald,M., Sandelin,A., Schneider,C., Semple,C.A., Setou,M., Shimada,K., Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tomita,M., Verardo,R., Wagner,L., Wahlstedt,C., Wang,Y., Watanabe,Y., Wells,C., Wilming,L.G., Wynshaw-Boris,A., Yanagisawa,M., Yang,I., Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A., Carninci,P., Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura,M., Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K., Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K., Ishii,Y., Itoh,M., Kigawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata,K., Shingawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S., Rogers,J., Birney,E. and Hayashizaki,Y.

TITLE
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

JOURNAL
Nature 420, 563-573 (2002)

PUBMED
12466851

COMMENT
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gscl.riken.jp, URL: http://genome.gsc.riken.jp/
Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hirozane,T., Imotani,K., Ishii,Y., Itoh,M., Kawai,J., Konno,H., Miyazaki,A., Murata,M., Nakamura,M., Nomura,K., Numazaki,R., Ohno,M., Sakai,K., Sakazume,N., Sasaki,D., Sato,K., Shibata,K., Shiraki,T., Tagami,M., Waki,K., Watahiki,A., Muramatsu,M. and Hayashizaki,Y. Direct Submission
Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Tissues were provided by Takashi Ishikawa (Department of Surgery 2 Yokohama City University 3-9 Fukuura, Kanazawa-ku, Yokohama 236-0004 Japan) whose assistance we gratefully acknowledge.
Please visit our web site (http://genome.gsc.riken.go.jp) for further details.

FEATURES
source
1. .382
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="I320011M17"
/cell_type="stroma cell"
/clone_lib="RIKEN full-length enriched, stroma cell"

ORIGIN
Query Match 93.6%; Score 23.4; DB 5; Length 382;
Best Local Similarity 96.0%; Pred. No. 5.8;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 AAAAAACGACACGAGCCCTCGG 25
|||||
Db 67 AAAAAACGACACGAGCCCTCGG 91
|||||

RESULT 41

BF235341
LOCUS
DEFINITION
602026734F1 NCI_CGAP_Li9 Mus musculus cDNA clone IMAGE:4161752 5', mRNA sequence.
BF235341
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Mus musculus (house mouse)

REFERENCE
1 (bases 1 to 383)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM9443 row: d column: 09
High quality sequence stop: 383.
Location/Qualifiers
1. .383
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4161752"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI CGAP Li9"
/note="Organ: liver; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.9 kb. Constructed by Life Technologies. Note: this is a NCI_CGAP Library."

FEATURES
source
1. .383
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4161752"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI CGAP Li9"
/note="Organ: liver; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.9 kb. Constructed by Life Technologies. Note: this is a NCI_CGAP Library."

ORIGIN

Query Match 93.6%; Score 23.4; DB 2; Length 383;
Best Local Similarity 96.0%; Pred. No. 5.8;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 AAAAAACGACACGAGCCCTCGG 25
|||||
Db 85 AAAAAACGACACGAGCCCTCGG 109
|||||

RESULT 42

BE197726
LOCUS
DEFINITION
us77c09.x1 NCI_CGAP_Mam4 Mus musculus cDNA clone IMAGE:3257200 3', mRNA
BE197726
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Mus musculus (house mouse)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Mus.
1 (bases 1 to 384)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Priscilla Furth Ph.D.

cdNA Library Preparation: Life Technologies, Inc.
 cdNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 image.lnl.gov/image/html/iresources.shtml

MG1:1068564

Seq primer: -40UP from Gibco
 High quality sequence stop: 383.

FEATURES

Location/Qualifiers
 1. .384
 /organism="Mus musculus"
 /mcl_type="mRNA"
 /strain="NMRI"
 /db_xref="taxon:10090"
 /clone="IMAGE:3257200"
 /tissue_type="tumor, gross tissue"
 /dev_stage="5 months"
 /lab_host="DH108"
 /clone_lib="NCI_CGAP Mam4"
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: Salt;
 Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.
 Library constructed by Life Technologies. Investigators
 providing samples: Lothar Hennighausen/Priscilla Furth.
 NIH Reference for transgenic model: Li et al., Cell Growth
 and Differentiation 7, 3-11 (1996)."

ORIGIN

Query Match 93.6%; Score 23.4; DB 2; Length 384;
 Best Local Similarity 96.0%; Pred. No. 5.8;
 Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAAAACGACACGCGAGCCCTCGG 25
 |||||
 Db 314 AAAAAACGACACTGCGAGCCCTCGG 290

RESULT 43

BY629024 384 bp mRNA linear EST 15-DEC-2002
 LOCUS BY629024 RIKEN full-length enriched, visual cortex Mus musculus
 DEFINITION cDNA clone K430028L21 3', mRNA sequence.

ACCESSION BY629024
 VERSION BY629024.1 GI:26964206
 KEYWORDS EST.

SOURCE

Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muroidae; Muridae; Murinae; Mus.

REFERENCE

1 (bases 1 to 384)
 Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
 Nikaado, I., Osato, N., Saito, R., Suzuki, H., Yamana, I.,
 Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,
 Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D. P., Bult, C.,
 Hume, D. A., Quackenbush, J., Schriml, L. M., Kanapin, A., Matsuda, H.,
 Batalov, S., Beisel, K. W., Blake, J. A., Bradt, D., Brusic, V.,
 Chothia, C., Corbani, L. E., Cousins, S., Dalla, E., Dragani, T. A.,
 Fletcher, C. F., Forrest, A., Frazer, K. S., Gaasterland, T.,
 Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S.,
 Gustincich, S., Hirokawa, N., Jackson, I. J., Jarvis, E. D., Kanai, A.,
 Kawaji, H., Kawasawa, Y., Kedzierski, R. M., King, B. L., Konagaya, A.,
 Kurochkin, I. V., Lee, Y., Lenhard, B., Lyons, P. A., Maglott, D. R.,
 Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T.,
 Numata, K., Okido, T., Pavan, W. J., Pertea, G., Pesole, G.,
 Petrovsky, N., Pillai, R., Pontius, J. U., Qi, D., Ramachandran, S.,
 Ravasi, T., Reed, J. C., Reed, D. J., Reid, J., Ring, B. Z., Ringwald, M.,
 Sandelin, A., Schneider, C., Semple, C. A., Setou, M., Shimada, K.,
 Sultana, R., Takenaka, Y., Taylor, M. S., Teasdale, R. D., Tomita, M.,
 Varardo, R., Wagner, L., Wallestedt, C., Wang, Y., Watanabe, Y.,
 Wells, C., Wilming, L. G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I.,
 Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P.,

Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M.,
 Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,
 Aizawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y.,
 Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K.,
 Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E. S.,
 Rogers, J., Birney, E. and Hayashizaki, Y.
 Analysis of the mouse transcriptome based on functional annotation
 of 60,770 full-length cDNAs

TITLE

JOURNAL

PUBLISHED

COMMENT

12466851
 Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/
 Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S.,
 Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H.,
 Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R.,
 Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K.,
 Shiraki, T., Tagami, M., Waki, K., Watanabe, A., Muramatsu, M. and
 Hayashizaki, Y. Direct Submission
 Computational Analysis of Full-Length Mouse cDNAs Compared with
 Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new
 genes. Genome Res. 10 (10), 1617-1630 (2000)
 RIKEN integrated sequence analysis (RISA) system-384-format
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.
 10 (11), 1757-1771 (2000)
 Computer-based methods for the mouse full-length cDNA
 encyclopedia: real-time sequence clustering for construction of a
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 cDNA library was prepared and sequenced in Mouse Genome
 Encyclopedia Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.
 Division of Experimental Animal Research in Riken contributed to
 prepare mouse tissues.
 Tissues were provided by Michela Pagliolini and Takao K. Hensch (
 Laboratory for Neuronal Circuit Development Brain Science Institute
 RIKEN 2-1 Hiroawa, Wako-shi, Saitama 351-0198 Japan) whose
 assistance we gratefully acknowledge.
 Please visit our web site (http://genome.gsc.riken.go.jp) for
 further details.

FEATURES

source

Location/Qualifiers
 1. .384
 /organism="Mus musculus"
 /mcl_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="K430028L21"
 /tissue_type="visual cortex"
 /clone_lib="RIKEN full-length enriched, visual cortex"

ORIGIN

Query Match 93.6%; Score 23.4; DB 5; Length 384;
 Best Local Similarity 96.0%; Pred. No. 5.8;
 Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAAAACGACACGCGAGCCCTCGG 25

|||||
 Db 68 AAAAAACGACACTGCGAGCCCTCGG 92

RESULT 44

BY699498 384 bp mRNA linear EST 16-DEC-2002
 LOCUS BY699498 RIKEN full-length enriched, osteoclast-like cell Mus
 DEFINITION musculus cDNA clone 1420117G23 3', mRNA sequence.
 ACCESSION BY699498
 VERSION BY699498.1 GI:27110532

KEYWORDS	EST.	
SOURCE	Mus musculus (house mouse)	
ORGANISM	Mus musculus	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridea; Muridae; Murinae; Mus.	
AUTHORS	1 (bases 1 to 384) Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nikaido, I., Oato, N., Saito, R., Suzuki, H., Yamanaka, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C., Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H., Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusci, V., Chothia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A., Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W.J., Perte, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S., Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M., Sandelin, A., Schneider, C., Semple, C.A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavalan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kigawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shingawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S., Rogers, J., Birney, E. and Hayashizaki, Y.	
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs	
JOURNAL	Nature 420, 563-573 (2002)	
FURNISHED	12466851	
COMMENT	Contact: Yoshihide Hayashizaki Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute The Institute of Physical and Chemical Research (RIKEN) 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan Tel: 81-45-503-9222 Fax: 81-45-503-9216 Email: genome.res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/ Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001) Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000) RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000) Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001) cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Tissues were provided by Takashi Ishikawa (Department of Surgery 236-0004 Japan) whose assistance we gratefully acknowledge. Please visit our web site (http://genome.gsc.riken.go.jp) for further details.	
FEATURES	Location/Qualifiers 1..384 /organism="Mus musculus" /mol_type="mRNA" /strain="CS7BL/6J" /db_xref="taxon:10090" /clone="I420117G23" /cell_type="osteoclast-like cell" /clone_lib="RIKEN full-length enriched, osteoclast-like cell"	
ORIGIN	Query Match 93.6%; Score 23.4; DB 5; Length 384; Best Local Similarity 96.0%; Pred. No. 5.8; Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0; QY 1 AAAAAACGACAGCGAGCCCTCGG 25 Db 69 AAAAAACGACACTGCGAGCCCTCGG 93 RESULT 45 BF320359/c LOCUS DEFINITION ACCESSION VERSION KEYWORDS SOURCE ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridea; Muridae; Murinae; Mus. 1 (bases 1 to 386) NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap. National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index Unpublished (1997) Other ESTs: uz52h08.y1 Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nih.gov Tissue Procurement: Jeffrey Green M.D. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL) DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at: image.llnl.gov/image/html/iresources.shtml MGI:1433503 High quality sequence stop: 385. Location/Qualifiers 1..386 /organism="Mus musculus" /mol_type="mRNA" /strain="FVB/N" /db_xref="taxon:10090" /clone="IMAGE:3672735" /sex="female, virgin" /tissue_type="infiltrating ductal carcinoma" /dev_stage="5 months" /lab_host="DH10B" /clone_lib="NCI_CGAP_Mam6" /note="Organ: mammary; Vector: pCMV-SPORT6; Site: 1; Salt; Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt. Library constructed by Life Technologies. Investigator providing samples: Jeffrey Green, M.D., NIH"	
ORIGIN	Query Match 93.6%; Score 23.4; DB 2; Length 386; Best Local Similarity 96.0%; Pred. No. 5.8; Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Tissues were provided by Michela Fagiolini and Takao K. Hensch (Laboratory for Neuronal Circuit Development Brain Science Institute RIKEN 2-1 Hirose, Wako-shi, Saitama 351-0198 Japan) whose assistance we gratefully acknowledge.

Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES
Location/Qualifiers
1..386
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CS7BL/6J"
/db_xref="taxon:10090"
/clone="K430348K09"
/tissue type="visual cortex"
/clone_lib="RIKEN full-length enriched, visual cortex"

ORIGIN
Query Match 93.6%; Score 23.4; DB 5; Length 386;
Best Local Similarity 96.0%; Pred. No. 5.8;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
1 AAAAAACGACACGAGCCCTCGG 25
|||||
78 AAAAAACGACACTGCGCCCTCGG 102

RESULT 47
BY603881
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS

BY603881 RIKEN full-length enriched, visual cortex Mus musculus
cDNA clone K230028E20 3', mRNA sequence.
BY603881
BY603881.1 GI:26939063
EST.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridea; Muridae; Murinae; Mus.
1 (bases 1 to 387)
Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nikaide, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C., Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H., Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusic, V., Chothia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A., Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A., Kawajiri, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W.J., Pertea, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S., Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M., Sandelin, A., Schneider, C., Semple, C.A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M., Verardo, R., Wagner, L.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Inotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K.,

1 AAAAAACGACACGAGCCCTCGG 25
|||||
319 AAAAAACGACACTGCGCCCTCGG 295

RESULT 46
BY642036
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS

BY642036 RIKEN full-length enriched, visual cortex Mus musculus
cDNA clone K430348K09 3', mRNA sequence.
BY642036
BY642036.1 GI:26977218
EST.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridea; Muridae; Murinae; Mus.
1 (bases 1 to 386)
Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nikaide, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C., Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H., Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusic, V., Chothia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A., Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A., Kawajiri, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W.J., Pertea, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S., Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M., Sandelin, A., Schneider, C., Semple, C.A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M., Verardo, R., Wagner, L.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Inotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Rogers, J., Birney, E. and Hayashizaki, Y.
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
12466851
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.jp, URL: <http://genome.gsc.riken.jp/>
Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hirozane, T., Inotani, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Sakai, M., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission
Computational Analysis of Full-length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771. (2000)

Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S., Rogers,J., Birney,E. and Hayashizaki,Y.
 Analysis of the mouse transcriptome based on functional annotation
 of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)
 1246851

TITLE

JOURNAL
PUBMED

COMMENT

Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center(GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216

Email: genome-res@gsr.riken.jp, URL:http://genome.gsc.riken.jp/
 Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S.,
 Hirozane,T., Imotani,K., Ishii,Y., Itoh,M., Kawai,J., Konno,H.,
 Miyazaki,A., Murata,M., Nakamura,M., Nomura,K., Numazaki,R.,
 Ohno,M., Sakai,K., Sakazume,N., Sasaki,D., Sato,K., Shibata,K.,
 Shiraki,T., Tagami,M., Waki,K., Watahiki,A., Muramatsu,M. and
 Hayashizaki,Y. Direct Submission
 Computational Analysis of Full-Length Mouse cDNAs Compared with
 Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new
 genes. Genome Res. 10 (10), 1617-1630 (2000)
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.
 10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA
 encyclopedia: real-time sequence clustering for construction of a
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 cDNA library was prepared and sequenced in Mouse Genome
 Encyclopedia Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.
 Division of Experimental Animal Research in Riken contributed to
 prepare mouse tissues.

Tissues were provided by Michela Fagioli and Takao K. Hensch (Laboratory for Neuronal Circuit Development Brain Science Institute RIKEN 2-1 Hiroawa,Wako-shi,Saitama 351-0198 Japan) whose assistance we gratefully acknowledge.

Please visit our web site (http://genome.gsc.riken.go.jp) for further details.

FEATURES
source

```
1. .387
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="K230028E20"
/tissue_type="visual cortex"
/clone_lib="RIKEN full-length enriched, visual cortex"
```

ORIGIN

Query Match . 93.6%; Score 23.4; DB 5; Length 387;
 Best Local Similarity 96.0%; Pred. No. 5.8;
 Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAAAACGACAGCGAGCCTCGG 25

```
|||||
|||||
|||||
```

Db

71 AAAAAACGACATCGAGCCTCGG 95

RESULT 48

BY630813

LOCUS

BY630813 RIKEN full-length enriched, visual cortex Mus musculus
 cDNA clone K430038M12 3', mRNA sequence.

ACCESSION

BY630813

VERSION

BY630813.1

KEYWORDS

EST.

SOURCE

Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE

AUTHORS

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muroidea; Muridae; Murinae; Mus.

1 (bases 1 to 387)

Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S.,
 Nikaido,I., Oeato,N., Saito,R., Suzuki,H., Yamataka,I.,
 Kiyosawa,H., Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A.,
 Schonbach,C., Gojibori,T., Baldarelli,R., Hill,D.P., Bult,C.,
 Hume,D.A., Quackenbush,J., Schriml,L.M., Kanapin,A., Matsuoka,H.,
 Batalov,S., Beisel,K.W., Blake,J.A., Bradt,D., Brusic,V.,
 Chothia,C., Corbani,L.E., Cousins,S., Della,E., Dragani,T.A.,
 Fletcher,C.F., Forrest,A., Frazer,K.S., Gasteland,I.,
 Gariboldi,M., Gissi,C., Godzik,A., Gough,J., Grimmond,S.,
 Gustincich,S., Hirokawa,N., Jackson,I.J., Jarvis,E.D., Kanai,A.,
 Kawai,H., Kawasawa,Y., Kedzierski,R.M., King,B.L., Konagaya,A.,
 Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R.,
 Maltais,L., Marchionni,L., McKenzie,L., Miki,H., Nagashima,T.,
 Numata,K., Okido,T., Pavan,W.J., Pertea,G., Pesole,G.,
 Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramachandran,S.,
 Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring,B.Z., Ringwald,M.,
 Sandelin,A., Schneider,C., Semple,C.A., Setou,M., Shimada,K.,
 Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tomita,M.,
 Verardo,R., Wagner,L., Wahlestedt,C., Wang,Y., Watanabe,Y.,
 Wells,C., Wilming,L.G., Wyshaw-Boris,A., Yanagisawa,M., Yang,I.,
 Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A., Carninci,P.,
 Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura,M.,
 Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K.,
 Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K., Ishii,Y.,
 Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata,K.,
 Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S.,
 Rogers,J., Birney,E. and Hayashizaki,Y.

Analysis of the mouse transcriptome based on functional annotation
 of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)
 1246851

TITLE

JOURNAL

PUBMED

COMMENT

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic

Sciences Center(GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-res@gsr.riken.jp, URL:http://genome.gsc.riken.jp/

Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S.,

Hirozane,T., Imotani,K., Ishii,Y., Itoh,M., Kawai,J., Konno,H.,

Miyazaki,A., Murata,M., Nakamura,M., Nomura,K., Numazaki,R.,

Ohno,M., Sakai,K., Sakazume,N., Sasaki,D., Sato,K., Shibata,K.,

Shiraki,T., Tagami,M., Waki,K., Watahiki,A., Muramatsu,M. and

Hayashizaki,Y. Direct Submission

Computational Analysis of Full-Length Mouse cDNAs Compared with

Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new

genes. Genome Res. 10 (10), 1617-1630 (2000)

RIKEN integrated sequence analysis (RISA) system--384-format

sequencing pipeline with 384 multicapillary sequencer. Genome Res.

10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA

encyclopedia: real-time sequence clustering for construction of a

nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

cDNA library was prepared and sequenced in Mouse Genome

Encyclopedia Project of Genome Exploration Research Group in Riken

Genomic Sciences Center and Genome Science Laboratory in RIKEN.

Division of Experimental Animal Research in Riken contributed to

prepare mouse tissues.

Tissues were provided by Michela Fagioli and Takao K. Hensch (

Laboratory for Neuronal Circuit Development Brain Science Institute

RIKEN 2-1 Hiroawa,Wako-shi,Saitama 351-0198 Japan) whose

assistance we gratefully acknowledge.

Please visit our web site (http://genome.gsc.riken.go.jp) for

further details.

Location/Qualifiers

1. .387

/organism="Mus musculus"

```

/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="K430038M12"
/tissue_type="visual cortex"
/clone_lib="RIKEN full-length enriched, visual cortex"

ORIGIN

Query Match      93.6%; Score 23.4; DB 5; Length 387;
Best Local Similarity 96.0%; Pred. No. 5.8;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAAAACGACACGAGCGCCCTCGG 25
    |||||
Db 74 AAAAAACGACACGAGCGCCCTCGG 98

RESULT 49
BY660055      387 bp      mRNA      linear      EST 16-DEC-2002
LOCUS      BY660055
DEFINITION      BY660055 RIKEN full-length enriched, visual cortex Mus musculus
                  cDNA clone K530360P19 3', mRNA sequence.
ACCESSION      BY660055
VERSION      BY660055.1 GI:27028389
KEYWORDS      Mus musculus (house mouse)
SOURCE      Mus musculus
ORGANISM      Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE      1 (bases 1 to 387)
AUTHORS      Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
              Nikaado I., Otsu N., Saito R., Suzuki H., Yamanaka I.,
              Kiyosawa H., Yagi K., Tomaru Y., Hasegawa Y., Nogami A.,
              Schonbach C., Gojobori T., Baldarelli R., Hill D.P., Bult C.,
              Hume D.A., Quackenbush J., Schriml L.M., Kanapin A., Matsuda H.,
              Bataillon S., Beisel K.W., Blake J.A., Bradt D., Brusic V.,
              Chothia C., Corbani L.E., Cousins S., Dalla E., Dragani T.A.,
              Fletcher C.F., Forrest A., Frazer K.S., Gaasterland T.,
              Gariboldi M., Gissi C., Godzik A., Gough J., Grimmond S.,
              Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D., Kanai A.,
              Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L., Konagaya A.,
              Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A., Maglott D.R.,
              Maltais L., Marchionni L., McKenzie L., Miki H., Nagashima T.,
              Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
              Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
              Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
              Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
              Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
              Varado R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y.,
              Wells C., Wilming L.G., Wyszynski B., Yang Y., Yanagisawa M., Yang I.,
              Yang L., Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P.,
              Hayatsu N., Hirozane-Kishikawa T., Konno H., Nakamura M.,
              Sakazume N., Sato K., Shiraki T., Waki K., Kawai J., Aizawa K.,
              Arakawa T., Fukuda S., Hara A., Hashizume W., Imotani K., Ishii Y.,
              Itoh M., Kagawa I., Miyazaki A., Sakai K., Sasaki D., Shibata K.,
              Shingawa A., Yasunishi A., Yoshino M., Waterston R., Lander E.S.,
              Rogers J., Birney E. and Hayashizaki Y.
              Analysis of the mouse transcriptome based on functional annotation
              of 60,770 full-length cDNAs
              Nature 420, 563-573 (2002)
              12466851
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Saitama-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@sc.riken.jp, URL: http://genome.gsc.riken.jp/
Aizawa K., Akimura T., Arakawa T., Carninci P., Fukuda S.,
Hirozane T., Imotani K., Ishii Y., Itoh M., Kawai J., Konno H.,
Miyazaki A., Murata M., Nakamura M., Nomura K., Numazaki R.,

```

```

Ohno M., Sakai K., Sakazume N., Sasaki D., Sato K., Shibata K.,
Shiraki T., Tagami M., Waki K., Watahiki A., Muramatsu M. and
Hayashizaki Y. Direct Submission
Computational Analysis of Full-length Mouse cDNAs Compared with
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Tissues were provided by Michela Fagiolini and Takao K. Hensch (
Laboratory for Neuronal Circuit Development Brain Science Institute
RIKEN 2-1 Hirosewa, Wako-shi, Saitama 351-0198 Japan ) whose
assistance we gratefully acknowledge.
Please visit our web site (http://genome.gsc.riken.go.jp) for
further details.

FEATURES             Location/Qualifiers
     1..387
         /organism="Mus musculus"
         /mol_type="mRNA"
         /strain="C57BL/6J"
         /db_xref="taxon:10090"
         /clone="K530360P19"
         /tissue_type="visual cortex"
         /clone_lib="RIKEN full-length enriched, visual cortex"

ORIGIN

Query Match      93.6%; Score 23.4; DB 5; Length 387;
Best Local Similarity 96.0%; Pred. No. 5.8;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAAAACGACACGAGCGCCCTCGG 25
    |||||
Db 72 AAAAAACGACACGAGCGCCCTCGG 96

RESULT 50
BY660055      388 bp      mRNA      linear      EST 15-OCT-2001
LOCUS      BY660055
DEFINITION      BY660055 RIKEN full-length enriched, 6 days neonate spleen Mus
                  musculus cDNA clone F430015P12 3', mRNA sequence.
ACCESSION      BY660055
VERSION      BY660055.1 GI:16136986
KEYWORDS      Mus musculus (house mouse)
SOURCE      Mus musculus
ORGANISM      Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE      1 (bases 1 to 388)
AUTHORS      Akimura T., Arakawa T., Carninci P., Furuno M., Hanagaki T.,
              Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T., Imotani K.,
              Ishii Y., Ito M., Kawai J., Kojima Y., Konno H., Kouda M.,
              Matsuyama T., Nakamura M., Nishi K., Nomura K., Numasaki R.,
              Okazaki Y., Okido T., Saito R., Sakai C., Sakai K., Sakazume N.,
              Sasaki D., Sato K., Shibata K., Shinagawa A., Shiraki T.,
              Sogabe Y., Suzuki H., Tagawa A., Takahashi F., Takaku-Akahira S.,
              Tanaka T., Tomaru A., Toya T., Watahiki A., Yasunishi A.,
              Muramatsu M. and Hayashizaki Y.
              RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura T., et al.
              2001)
              Unpublished (2001)
              Contact: Yoshihide Hayashizaki
              Laboratory for Genome Exploration Research Group, RIKEN Genomic

```


GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.
OM nucleic - nucleic search, using sw model
Run on: February 3, 2006, 11:35:31 ; Search time 80.6667 Seconds
(without alignments)
550.897 Million cell updates/sec

Title: US-10-719-900-15
Perfect score: 25
Sequence: 1 aaaaacacacagcagccctcgg 25
Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues
Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 150 summaries

Database : Issued Patents NA: *
1: /cgn2_6/ptodata/1/ina/1 COMB.seq: *
2: /cgn2_6/ptodata/1/ina/5 COMB.seq: *
3: /cgn2_6/ptodata/1/ina/6A COMB.seq: *
4: /cgn2_6/ptodata/1/ina/6B COMB.seq: *
5: /cgn2_6/ptodata/1/ina/H COMB.seq: *
6: /cgn2_6/ptodata/1/ina/PCTUS COMB.seq: *
7: /cgn2_6/ptodata/1/ina/PP COMB.seq: *
8: /cgn2_6/ptodata/1/ina/RE COMB.seq: *
9: /cgn2_6/ptodata/1/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	17.6	70.4	32881	3	US-09-949-016-16048
C 2	17	68.0	601	3	US-09-949-016-59715
C 3	17	68.0	603	3	US-09-489-039A-6163
C 4	17	68.0	19857	3	US-09-949-016-13114
C 5	17	68.0	19998	3	US-09-949-016-12657
C 6	17	68.0	36938	3	US-09-949-016-13484
C 7	16.6	66.4	255	3	US-09-248-796A-11154
C 8	16.6	66.4	601	3	US-09-949-002-1163
C 9	16.6	66.4	601	3	US-09-949-002-7259
C 10	16.6	66.4	2025	3	US-08-991-677-7
C 11	16.6	66.4	4493	3	US-09-949-016-14034
C 12	16.6	66.4	40325	3	US-09-949-002-580
C 13	16.6	66.4	40325	3	US-09-949-002-771
C 14	16.6	66.4	43717	3	US-09-949-016-16821
C 15	16.2	64.8	452	3	US-09-300-958A-40
C 16	16.2	64.8	1409	9	5242798-6
C 17	16.2	64.8	2826	3	US-09-902-540-3646
C 18	16.2	64.8	16563	3	US-09-902-540-2000
C 19	16.2	64.8	16844	3	US-09-902-540-1176
C 20	16.2	64.8	28493	3	US-09-902-540-1241
C 21	16.2	64.8	41322	3	US-10-024-396-13
C 22	16.2	64.8	90472	3	US-09-949-016-14038
C 23	16.2	64.8	158735	3	US-09-949-016-11989
C 24	16.2	64.8	158735	3	US-09-949-016-17130

C 25	16	64.0	184	3	US-09-513-999C-16876	Sequence 16876, A
C 26	16	64.0	531	3	US-09-533-559-2241	Sequence 2241, Ap
C 27	16	64.0	601	3	US-09-949-016-27169	Sequence 27169, A
C 28	16	64.0	601	3	US-09-949-016-178692	Sequence 178692, A
C 29	16	64.0	601	3	US-09-949-016-178693	Sequence 178693, A
C 30	16	64.0	601	3	US-09-949-016-182367	Sequence 182367, A
C 31	16	64.0	1347	3	US-09-543-681A-2689	Sequence 2689, Ap
C 32	16	64.0	1600	3	US-09-949-016-651	Sequence 651, App
C 33	16	64.0	13595	3	US-09-949-016-12529	Sequence 12529, A
C 34	16	64.0	13970	3	US-09-949-016-16690	Sequence 16690, A
C 35	16	64.0	27968	3	US-09-949-016-15191	Sequence 15191, A
C 36	16	64.0	27968	3	US-09-949-016-15192	Sequence 15192, A
C 37	16	64.0	34372	3	US-09-949-016-13098	Sequence 13098, A
C 38	16	64.0	34875	3	US-09-949-016-13099	Sequence 13099, A
C 39	16	64.0	37861	3	US-09-949-016-16872	Sequence 16872, A
C 40	16	64.0	97195	3	US-09-949-016-12212	Sequence 12212, A
C 41	16	64.0	97196	3	US-09-949-016-16971	Sequence 16971, A
C 42	15.8	63.2	571	3	US-08-779-764A-13	Sequence 13, Appl
C 43	15.8	63.2	571	3	US-09-563-456-13	Sequence 13, Appl
C 44	15.8	63.2	1692	3	US-09-252-991A-7872	Sequence 7872, Ap
C 45	15.8	63.2	6755	3	US-08-931-999-4	Sequence 4, Appl
C 46	15.8	63.2	8910	3	US-08-779-764A-1	Sequence 1, Appl
C 47	15.8	63.2	8910	3	US-09-563-456-1	Sequence 1, Appl
C 48	15.8	63.2	72662	3	US-09-949-016-17073	Sequence 17073, A
C 49	15.8	63.2	131254	3	US-09-949-016-13734	Sequence 13734, A
C 50	15.8	63.2	135030	3	US-09-949-016-14896	Sequence 14896, A
C 51	15.6	62.4	111	3	US-09-702-705-1556	Sequence 1556, Ap
C 52	15.6	62.4	111	3	US-09-736-457-1556	Sequence 1556, Ap
C 53	15.6	62.4	111	3	US-09-614-1248-1556	Sequence 1556, Ap
C 54	15.6	62.4	111	3	US-09-671-325-1556	Sequence 1556, Ap
C 55	15.6	62.4	111	3	US-09-658-824-1556	Sequence 1556, Ap
C 56	15.6	62.4	111	3	US-10-017-754-1556	Sequence 1556, Ap
C 57	15.6	62.4	111	3	US-09-651-563-1556	Sequence 1556, Ap
C 58	15.6	62.4	258	3	US-09-543-681A-1487	Sequence 1487, Ap
C 59	15.6	62.4	309	3	US-09-513-999C-32302	Sequence 32302, A
C 60	15.6	62.4	315	3	US-09-621-976-17895	Sequence 17895, A
C 61	15.6	62.4	414	3	US-08-905-223-191	Sequence 191, App
C 62	15.6	62.4	426	3	US-09-621-976-17892	Sequence 17892, A
C 63	15.6	62.4	466	3	US-09-513-999C-886	Sequence 886, App
C 64	15.6	62.4	601	3	US-09-949-016-78047	Sequence 78047, A
C 65	15.6	62.4	601	3	US-09-949-016-139389	Sequence 139389, A
C 66	15.6	62.4	601	3	US-09-949-016-139390	Sequence 139390, A
C 67	15.6	62.4	601	3	US-09-949-016-139391	Sequence 139391, A
C 68	15.6	62.4	601	3	US-09-949-016-139392	Sequence 139392, A
C 69	15.6	62.4	601	2	US-09-949-016-169403	Sequence 169403, A
C 70	15.6	62.4	610	2	US-08-928-926A-2	Sequence 2, Appl
C 71	15.6	62.4	610	3	US-09-212-149-2	Sequence 2, Appl
C 72	15.6	62.4	666	3	US-09-543-681A-147	Sequence 147, App
C 73	15.6	62.4	954	3	US-09-252-991A-15575	Sequence 15575, A
C 74	15.6	62.4	1098	3	US-09-774-528-183	Sequence 183, App
C 75	15.6	62.4	1098	3	US-10-120-988-183	Sequence 183, App
C 76	15.6	62.4	1116	3	US-09-252-991A-15677	Sequence 15677, A
C 77	15.6	62.4	5999	3	US-09-949-016-13336	Sequence 13336, A
C 78	15.6	62.4	22294	3	US-09-949-016-14020	Sequence 14020, A
C 79	15.6	62.4	45432	3	US-09-949-002-719	Sequence 719, App
C 80	15.6	62.4	80246	3	US-09-078-294-4	Sequence 4, Appl
C 81	15.6	62.4	80595	3	US-09-078-294-3	Sequence 3, Appl
C 82	15.6	62.4	117001	3	US-09-949-016-15684	Sequence 15684, A
C 83	15.6	62.4	127771	3	US-09-949-016-14982	Sequence 14982, A
C 84	15.6	62.4	177293	3	US-09-949-016-16313	Sequence 16313, A
C 85	15.6	62.4	258775	3	US-09-949-016-16435	Sequence 16435, A
C 86	15.4	61.6	232	3	US-09-513-999C-33162	Sequence 33162, A
C 87	15.4	61.6	375	3	US-09-270-767-15003	Sequence 15003, A
C 88	15.4	61.6	462	3	US-09-270-767-6285	Sequence 6285, Ap
C 89	15.4	61.6	462	3	US-09-270-767-21567	Sequence 21567, A
C 90	15.4	61.6	488	3	US-09-270-767-4242	Sequence 4242, Ap
C 91	15.4	61.6	488	3	US-09-270-767-19524	Sequence 19524, A
C 92	15.4	61.6	494	3	US-09-533-559-5232	Sequence 5232, Ap
C 93	15.4	61.6	601	3	US-09-949-016-42914	Sequence 42914, A
C 94	15.4	61.6	601	3	US-09-949-016-42915	Sequence 42915, A
C 95	15.4	61.6	601	3	US-09-949-016-42916	Sequence 42916, A
C 96	15.4	61.6	601	3	US-09-949-016-59081	Sequence 59081, A
C 97	15.4	61.6	601	3	US-09-949-016-59082	Sequence 59082, A

Sequence 109227, Sequence 119333, Sequence 119334, Sequence 119334, Sequence 39, Appl Sequence 36, Appl Sequence 2047, Ap Sequence 6100, Ap Sequence 711, App Sequence 450, App Sequence 457, App Sequence 8, Appl Sequence 16, Appl Sequence 1, Appl Sequence 1, Appl Sequence 12552, A Sequence 15435, A Sequence 65, Appl Sequence 13240, A Sequence 15906, A Sequence 836, App Sequence 17566, A Sequence 12752, A Sequence 15115, A Sequence 3, Appl Sequence 3, Appl Sequence 3, Appl Sequence 13408, A Sequence 581, App Sequence 17278, A Sequence 11900, A Sequence 13834, A Sequence 13453, A Sequence 15051, A Sequence 12859, A Sequence 14757, A Sequence 16089, A Sequence 577, App Sequence 705, App Sequence 12980, A Sequence 11842, A Sequence 15753, A Sequence 31, Appl Sequence 34, Appl Sequence 14033, A Sequence 12777, A Sequence 6856, Ap Sequence 22138, A Sequence 1849, Ap

US-09-949-016-109227 3 US-09-949-016-119333 3 US-09-949-016-119334 3 US-08-545-809A-39 3 US-09-515-697-39 3 US-10-144-929-36 3 US-09-543-681A-2047 3 US-09-533-559-6100 3 US-09-540-236-711 3 US-09-489-039A-2550 3 US-09-489-039A-457 3 US-08-267-803B-8 3 US-09-041-886-16 3 US-07-920-281C-1 3 US-08-466-277-1 3 US-09-688-842-1 3 US-09-949-016-12552 3 US-09-949-016-15435 3 US-09-719-554-65 3 US-09-949-016-13240 3 US-09-596-002-12 3 US-09-949-016-15906 3 US-09-949-002-836 3 US-09-949-016-17566 3 US-09-949-016-12752 3 US-09-949-016-15115 3 US-09-735-934A-3 3 US-10-060-332-3 3 US-10-339-657-3 3 US-10-885-879-3 3 US-09-949-016-13408 3 US-09-949-002-581 3 US-09-949-016-17278 3 US-09-949-016-17207 3 US-09-949-016-11900 3 US-09-949-016-13834 3 US-09-949-016-13453 3 US-09-949-016-15051 3 US-09-949-016-12859 3 US-09-949-016-14757 3 US-09-949-016-16089 3 US-09-949-002-577 3 US-09-949-002-705 3 US-09-949-016-12980 3 US-09-949-016-11842 3 US-09-949-016-15753 3 US-09-949-001-31 3 US-09-949-001-34 3 US-09-949-016-14033 3 US-09-949-016-12777 3 US-09-270-767-6856 3 US-09-270-767-22138 3 US-09-640-211A-1849 3

ALIGNMENTS

RESULT 1
US-09-949-016-16048/c
; Sequence 16048, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949, 016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03

Sequence 109227, Sequence 119333, Sequence 119334, Sequence 119334, Sequence 39, Appl Sequence 36, Appl Sequence 2047, Ap Sequence 6100, Ap Sequence 711, App Sequence 450, App Sequence 457, App Sequence 8, Appl Sequence 16, Appl Sequence 1, Appl Sequence 1, Appl Sequence 12552, A Sequence 15435, A Sequence 65, Appl Sequence 13240, A Sequence 15906, A Sequence 836, App Sequence 17566, A Sequence 12752, A Sequence 15115, A Sequence 3, Appl Sequence 3, Appl Sequence 3, Appl Sequence 13408, A Sequence 581, App Sequence 17278, A Sequence 11900, A Sequence 13834, A Sequence 13453, A Sequence 15051, A Sequence 12859, A Sequence 14757, A Sequence 16089, A Sequence 577, App Sequence 705, App Sequence 12980, A Sequence 11842, A Sequence 15753, A Sequence 31, Appl Sequence 34, Appl Sequence 14033, A Sequence 12777, A Sequence 6856, Ap Sequence 22138, A Sequence 1849, Ap

US-09-949-016-16048/c
; Sequence 16048, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949, 016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 59715
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-59715

Query Match 68.0%; Score 17; DB 3; Length 601;
Best Local Similarity 80.0%; Pred. No. 2.3e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAAACGACACAGCGAGCCCTCGG 25
|||||
DB 476 AAGATCCACACAGTGAGCCCTCG 500

RESULT 3
US-09-489-039A-6163/c
; Sequence 6163, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 6163
; LENGTH: 603

; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-6163

Query Match 68.0%; Score 17; DB 3; Length 603;
Best Local Similarity 80.0%; Pred. No. 2.3e+02;
Matches 20; Conservative 0; Mismatches 5; Indels

QY

1 AAAAAACGACACAGCGAGCCCTCGG 25
||| ||| ||| ||| ||| ||| ||| |||
Dp

164 AAATAGCGAAACAGCGTTTCCCTCGG 140

```

RESULT 4
US-09-949-016-13114
; Sequence 13114, Application US/09949016
; Patent No. 681239
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13114
; LENGTH: 19857
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13114

```

Query Match	68.0%;	Score 17;	DB 3;	Length 19857;
Best Local Similarity	80.0%;	Pred. No. 2.8e+02;		
Matches 20:	Conservative	0: Mismatches	5: Indels	0: Gaps
				0:

QY 1 AAAAAACGACACAGCGAGCCCTCGG 25
||||| ||||| ||||| |||||
DB 3315 AAAAAAAACACAGCCAGCCCAAGG 3339

```

RESULT 5
US-09-949-016-12657
; Sequence 12657, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12657
; LENGTH: 19998
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-12657

```

Query Match 68.0%; Score 17; DB 3; Length 19998;

Best Local Similarity	80.0%	Pred. No. 2.8e+02;	
Matches	20: Conservative	0; Mismatches	5; Indels
		0; Gaps	0;

Qy

1 AAAAAACGACACAGCGAGCCCTCG 25
||||| ||||| ||||| ||||| |||||

D_b

3259 AAAAAAACACAGCCAGCCCCAAG 3283

```

RESULT 6
US-09-949-016-13484
; Sequence 13484, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13484
; LENGTH: 36938
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13484

```

Query Match 68.0%; Score 17; DB 3; Length 36938;
Best Local Similarity 80.0%; Pred. No. 2.9e+02;
Matches 20: Conservative 0; Mismatches 5; Indels

QY 1 AAAAAACGACACAGCGAGCCCTCGG 25
||| | | | | | | | | | | | |
Dp 21572 AAAGATCCACACAGTGAGCCCTCTG 21596

```

RESULT 7
US-09-248-796A-11154/C
, Sequence 11154, Application US/09248796A
, Patent No. 6747137
, GENERAL INFORMATION:
, APPLICANT: Keith Weinstock et al
, TITLE OF INVENTION: NUCLEIC ACID AND AMINO
, TITLE OF INVENTION: FOR DIAGNOSTICS AND T
, FILE REFERENCES: 107196.132
, CURRENT APPLICATION NUMBER: US/09/248,796A
, CURRENT FILING DATE: 1999-02-12
, PRIOR APPLICATION NUMBER: US 60/074,725
, PRIOR FILING DATE: 1998-02-13
, PRIOR APPLICATION NUMBER: US 60/096,409
, PRIOR FILING DATE: 1998-08-13
, NUMBER OF SEQ ID NOS: 28208
, SEQ ID NO 11154
, LENGTH: 255
, TYPE: DNA
, ORGANISM: Candida albicans
US-09-248-796A-11154

```

```

Query Match          66.4%; Score 16.6; DB 3; Length 255;
Best Local Similarity 82.6%; Pred. No. 3.3e+02;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

QY 1 AAAAAACGACACAGCGAGCCCTC 23
||||| ||| ||| ||| ||| |||
Db 250 AAAAAAGAGACATCGAGCCCCC 228

; SEQ ID NO 580
; LENGTH: 40325
; TYPE: DNA
; ORGANISM: Human
US-09-949-002-580

Query Match 66.4%; Score 16.6; DB 3; Length 40325;
Best Local Similarity 82.6%; Pred. No. 4.4e+02;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAACGACAGCGAGCCCTC 23
DB 24633 AAAAAAAGAGCGAGCCCTC 24611
||||| |

RESULT 13

US-09-949-002-771/c
; Sequence 771, Application US/09949002
; Patent No. 6900016

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION

; TITLE OF INVENTION: AND USES THEREOF

; FILE REFERENCE: CL000790

; CURRENT APPLICATION NUMBER: US/09/949,002

; CURRENT FILING DATE: 2000-01-28

; PRIOR APPLICATION NUMBER: 60/231,401

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 10823

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 771

; LENGTH: 40325

; TYPE: DNA

; ORGANISM: Human

US-09-949-002-771

Query Match 66.4%; Score 16.6; DB 3; Length 40325;
Best Local Similarity 82.6%; Pred. No. 4.4e+02;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAACGACAGCGAGCCCTC 23
DB 24633 AAAAAAAGAGCGAGCCCTC 24611
||||| |

RESULT 14

US-09-949-016-16821/c
; Sequence 16821, Application US/09949016
; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 16821

; LENGTH: 43717

; TYPE: DNA

; ORGANISM: Human

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (1)-(43717)

; OTHER INFORMATION: n = A,T,C or G

US-09-949-016-16821

Query Match 66.4%; Score 16.6; DB 3; Length 43717;
Best Local Similarity 82.6%; Pred. No. 4.4e+02;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 AAAAAACGACAGCGAGCCCTCGG 25
DB 19034 AAAAAACACTGCAGCCCTGGG 19012
||||| |

RESULT 15

US-09-300-958A-40

; Sequence 40, Application US/09300958A

; Patent No. 6495319

; GENERAL INFORMATION:

; APPLICANT: McClelland, Michael

; APPLICANT: Welsh, John

; APPLICANT: Trenkle, Thomas

; TITLE OF INVENTION: Reduced Complexity Nucleic Acid Targets and Methods of

; FILE REFERENCE: P-PH 3457

; CURRENT APPLICATION NUMBER: US/09/300,958A

; CURRENT FILING DATE: 1999-04-27

; PRIOR APPLICATION NUMBER: 60/083,331

; PRIOR FILING DATE: 1998-04-27

; PRIOR APPLICATION NUMBER: 60/098,070

; PRIOR FILING DATE: 1998-08-27

; PRIOR APPLICATION NUMBER: 60/118,624

; PRIOR FILING DATE: 1999-02-04

; NUMBER OF SEQ ID NOS: 85

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 40

; LENGTH: 452

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: unsure

; LOCATION: (33)-(34)

; NAME/KEY: unsure

; LOCATION: (59)

; NAME/KEY: unsure

; LOCATION: (82)

; NAME/KEY: unsure

; LOCATION: (112)

; NAME/KEY: unsure

; LOCATION: (126)

; NAME/KEY: unsure

; LOCATION: (164)

; NAME/KEY: unsure

; LOCATION: (184)

; NAME/KEY: unsure

; LOCATION: (225)

; NAME/KEY: unsure

; LOCATION: (244)

; NAME/KEY: unsure

; LOCATION: (253)

; NAME/KEY: unsure

; LOCATION: (272)

; NAME/KEY: unsure

; LOCATION: (307)

; NAME/KEY: unsure

; LOCATION: (316)

; NAME/KEY: unsure

; LOCATION: (329)

; NAME/KEY: unsure

; LOCATION: (335)

; NAME/KEY: unsure

; LOCATION: (381)

; NAME/KEY: unsure

; LOCATION: (396)

; NAME/KEY: unsure

; LOCATION: (417)


```
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 1241
; LENGTH: 28493
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(28493)
; OTHER INFORMATION: unsure at all n locations
US-09-902-540-1241

Query Match          64.8%; Score 16.2; DB 3; Length 28493;
Best Local Similarity 85.7%; Pred. No. 6.6e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      5 AACGACACGAGCGCCCTCGG 25
Db      17737 AACGATACAGCGAGCCACCG 17717

RESULT 21
US-10-024-396-13/c
; Sequence 13, Application US/10024396
; Patent No. 6852536
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Dobie
; TITLE OF INVENTION: ANTISENSE MODULATION OF CD36L1 EXPRESSION
; FILE REFERENCE: RTS-0339
; CURRENT APPLICATION NUMBER: US/10/024,396
; CURRENT FILING DATE: 2001-12-18
; NUMBER OF SEQ ID NOS: 91
; SEQ ID NO 13
; LENGTH: 41322
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: exon:intron junction
; LOCATION: (9980)...(9981)
; OTHER INFORMATION: exon 8:intron 8
; NAME/KEY: intron
; LOCATION: (17618)...(22472)
; OTHER INFORMATION: intron 9
; NAME/KEY: intron
; LOCATION: (22547)...(30283)
; OTHER INFORMATION: intron 10
; NAME/KEY: intron:exon junction
; LOCATION: (30283)...(30284)
; OTHER INFORMATION: intron 10:exon 11
; NAME/KEY: intron:exon junction
; LOCATION: (31237)...(31238)
; OTHER INFORMATION: intron 11:exon 12
; NAME/KEY: intron
; LOCATION: (31385)...(34929)
; OTHER INFORMATION: intron 12
; NAME/KEY: intron
; LOCATION: (35059)...(39154)
; OTHER INFORMATION: intron 13
; NAME/KEY: intron:exon junction
; LOCATION: (39154)...(39155)
; OTHER INFORMATION: intron 13:exon 14
US-10-024-396-13

Query Match          64.8%; Score 16.2; DB 3; Length 41322;
Best Local Similarity 85.7%; Pred. No. 6.7e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
QY      3 AAAACGACACGAGCGCCCTC 23
Db      7922 AACAGACACAGCAAGGCCTC 7902

RESULT 22
US-09-949-016-14038/c
; Sequence 14038, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14038
; LENGTH: 90472
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(90472)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14038

Query Match          64.8%; Score 16.2; DB 3; Length 90472;
Best Local Similarity 85.7%; Pred. No. 6.9e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      3 AAAACGACACGAGCGCCCTC 23
Db      56403 AACACGACACAGCAAGGCCTC 56383

RESULT 23
US-09-949-016-11989/c
; Sequence 11989, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11989
; LENGTH: 158735
; TYPE: DNA
; ORGANISM: Human
; OTHER INFORMATION:
US-09-949-016-11989

Query Match          64.8%; Score 16.2; DB 3; Length 158735;
Best Local Similarity 85.7%; Pred. No. 7.1e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 1 AAAAAACGACACGAGCGGCC 21
Db 55134 ACAAAGACACGACGAGGCC 55114

RESULT 24
US-09-949-016-17130/c
; Sequence 17130, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17130
; LENGTH: 158735
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17130

Query Match 64.8%; Score 16.2; DB 3; Length 158735;
Best Local Similarity 85.7%; Pred. No. 7.1e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAACGACACGAGCGGCC 21
Db 55134 ACAAAGACACGACGAGGCC 55114

RESULT 25
US-09-513-999C-16876/c
; Sequence 16876, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 16876
; LENGTH: 184
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-513-999C-16876

Query Match 64.0%; Score 16; DB 3; Length 184;
Best Local Similarity 79.2%; Pred. No. 6e+02;
Matches 19; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 AAAAAACGACACGAGCGGCCCTCGG 25
Db 102 AAAATCGTCACAGACCTCTCGG 79

RESULT 26
US-09-533-559-2241

; Sequence 2241, Application US/09533559
; Patent No. 6902887
; GENERAL INFORMATION:
; APPLICANT: Randy M. Berka
; APPLICANT: Michael W. Key
; APPLICANT: Jeffrey R. Shuster
; APPLICANT: Sakari Kauppinen
; APPLICANT: Ib Groth Clausen
; APPLICANT: Peter Bjarke Olsen
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; FILE REFERENCE: 5849.200-US
; CURRENT APPLICATION NUMBER: US/09/533,559
; CURRENT FILING DATE: 2000-03-22
; EARLIER APPLICATION NUMBER: 09/273,623
; EARLIER FILING DATE: 1999-03-22
; NUMBER OF SEQ ID NOS: 7860
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2241
; LENGTH: 531
; TYPE: DNA
; ORGANISM: Fusarium venenatum
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(531)
; OTHER INFORMATION: n = A,T,C or G
US-09-533-559-2241

Query Match 64.0%; Score 16; DB 3; Length 531;
Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 CGACACGAGCGCCCT 22
Db 217 CGACACGAGCGCCCT 232

RESULT 27
US-09-949-016-27169/c
; Sequence 27169, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27169
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-27169

Query Match 64.0%; Score 16; DB 3; Length 601;
Best Local Similarity 79.2%; Pred. No. 6.4e+02;
Matches 19; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 AAAAAACGACACGAGCGGCCCTCGG 25
Db 144 AAAAATGACACAGGCTGCACTCGG 121

RESULT 28
US-09-949-016-178692
```

```
; Sequence 178692, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 178692
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-178692

Query Match          64.0%; Score 16; DB 3; Length 601;
Best Local Similarity 79.2%; Pred. No. 6.4e+02;
Matches 19; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 AAAAAGCAGACGAGCGCCCTCGG 25
Db 36 ACACAGCAGACAGAGAGCCCGGG 59

RESULT 29
US-09-949-016-178693
; Sequence 178693, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 178693
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-178693

Query Match          64.0%; Score 16; DB 3; Length 601;
Best Local Similarity 79.2%; Pred. No. 6.4e+02;
Matches 19; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 AAAAAGCAGACGAGCGCCCTCGG 25
Db 536 ACACAGCAGACAGAGAGCCCGGG 559

RESULT 30
US-09-949-016-182367/c
; Sequence 182367, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
```

```
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 182367
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-182367

Query Match          64.0%; Score 16; DB 3; Length 601;
Best Local Similarity 79.2%; Pred. No. 6.4e+02;
Matches 19; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 AAAAAGCAGACGAGCGCCCTCGG 25
Db 144 AAAAATGACACAGAGCGTGCACCTCGG 121

RESULT 31
US-09-543-681A-2689/c
; Sequence 2689, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILI
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 2689
; LENGTH: 1347
; TYPE: DNA
; ORGANISM: Proteus mirabilis
; US-09-543-681A-2689

Query Match          64.0%; Score 16; DB 3; Length 1347;
Best Local Similarity 79.2%; Pred. No. 6.7e+02;
Matches 19; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 AAAAAGCAGACGAGCGCCCTCGG 25
Db 876 AATAATGCCACACAGAGAGCCCGGG 853

RESULT 32
US-09-949-016-651
; Sequence 651, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
```

```
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 651
; LENGTH: 1600
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-651

Query Match          64.0%; Score 16; DB 3; Length 1600;
Best Local Similarity 79.2%; Pred. No. 6.8e+02;
Matches 19; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 AAAAAACGACACGAGCGCCCTCGG 25
   ||||| ||||| ||||| ||||| |||||
Db 322 AGAAATGACACGCGAGCGCATCGG 345

RESULT 33
US-09-949-016-12529/c
; Sequence 12529, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12529
; LENGTH: 13595
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12529

Query Match          64.0%; Score 16; DB 3; Length 13595;
Best Local Similarity 79.2%; Pred. No. 7.8e+02;
Matches 19; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAAACGACACGAGCGCCCTCG 24
   ||||| ||||| ||||| ||||| |||||
Db 2973 AAATACCACACGACGAGCCCTTCG 2950

RESULT 34
US-09-949-016-16690/c
; Sequence 16690, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16690
; LENGTH: 13970
; TYPE: DNA
; ORGANISM: Human

; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 651
; LENGTH: 1600
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-651

Query Match          64.0%; Score 16; DB 3; Length 13970;
Best Local Similarity 79.2%; Pred. No. 7.8e+02;
Matches 19; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAAACGACACGAGCGCCCTCG 24
   ||||| ||||| ||||| ||||| |||||
Db 3018 AAATACCACACGACGAGCCCTTCG 2995

RESULT 35
US-09-949-016-15191/c
; Sequence 15191, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15191
; LENGTH: 27968
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(27968)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15191

Query Match          64.0%; Score 16; DB 3; Length 27968;
Best Local Similarity 79.2%; Pred. No. 8e+02;
Matches 19; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAAACGACACGAGCGCCCTCG 24
   ||||| ||||| ||||| ||||| |||||
Db 20778 AAAAAAGCGACGAGAGCGCCCTCG 20755

RESULT 36
US-09-949-016-15192/c
; Sequence 15192, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15192
; LENGTH: 27968
; TYPE: DNA
; ORGANISM: Human
```

```
;
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(27968)
; OTHER INFORMATION: n = A,T,C or G
; US-09-949-016-15192

Query Match      64.0%; Score 16; DB 3; Length 27968;
Best Local Similarity 79.2%; Pred. No. 8e+02;
Matches 19; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAAAGCAGCAGCGAGCCCTCG 24
   ||||| ||| ||||| |||||
Db 20778 AAAAAAGCAGCAGCGAGCCCTCG 20755

RESULT 37
US-09-949-016-13098/c
; Sequence 13098, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13098
; LENGTH: 34372
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-13098

Query Match      64.0%; Score 16; DB 3; Length 34372;
Best Local Similarity 79.2%; Pred. No. 8.1e+02;
Matches 19; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 AAAAAAGCAGCAGCGAGCCCTCG 25
   ||||| ||||| ||||| |||||
Db 3045 AAAAAAGCAGCAGCGAGCCCTCG 3022

RESULT 38
US-09-949-016-13099/c
; Sequence 13099, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13099
; LENGTH: 34875
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-13099

Query Match      64.0%; Score 16; DB 3; Length 34875;
Best Local Similarity 79.2%; Pred. No. 8.1e+02;
Matches 19; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 AAAAAAGCAGCAGCGAGCCCTCG 25
   ||||| ||||| ||||| |||||
Db 3045 AAAAAAGCAGCAGCGAGCCCTCG 3022

RESULT 39
US-09-949-016-16872/c
; Sequence 16872, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16872
; LENGTH: 37861
; TYPE: DNA
; ORGANISM: Human
; NAME/KEY: misc_feature
; LOCATION: (1)...(37861)
; OTHER INFORMATION: n = A,T,C or G
; US-09-949-016-16872

Query Match      64.0%; Score 16; DB 3; Length 37861;
Best Local Similarity 79.2%; Pred. No. 8.2e+02;
Matches 19; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 AAAAAAGCAGCAGCGAGCCCTCG 25
   ||||| ||||| ||||| |||||
Db 31151 ACACAGCAGCAGAGAGCCCGG 31128

RESULT 40
US-09-949-016-12212/c
; Sequence 12212, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12212
; LENGTH: 97195
; TYPE: DNA
; ORGANISM: Human
; NAME/KEY: misc_feature
; US-09-949-016-12212
```

```
; LOCATION: (1)...(97195)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12212

Query Match          64.0%; Score 16; DB 3; Length 97195;
Best Local Similarity 79.2%; Pred. No. 8.5e+02;
Matches 19; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 AAAAAGCACACGAGCGCCTCGG 25
   ||||| ||||| ||||| ||||| |||||
Db 31799 AAAAATGCACACAGGCTGCACTCGG 31776

RESULT 41
US-09-949-016-16971/c
; Sequence 16971, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16971
; LENGTH: 97196
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(97196)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16971

Query Match          64.0%; Score 16; DB 3; Length 97196;
Best Local Similarity 79.2%; Pred. No. 8.5e+02;
Matches 19; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 AAAAAGCACACGAGCGCCTCGG 25
   ||||| ||||| ||||| ||||| |||||
Db 31799 AAAAATGCACACAGGCTGCACTCGG 31776

RESULT 42
US-08-779-764A-13
; Sequence 13, Application US/08779764
; Patent No. 6057094
; GENERAL INFORMATION:
; APPLICANT: de la Torre, Juan C.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR SCREENING
; TITLE OF INVENTION: OF HUMAN BORNA DISEASE VIRUS
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: THE SCRIPPS RESEARCH INSTITUTE
; STREET: 10550 No. 6057094th Torrey Pines Road, TPC-8
; CITY: La Jolla
; STATE: California
; COUNTRY: United States
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/779,764
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitting, Thomas
; REGISTRATION NUMBER: 34,163
; REFERENCE/DOCKET NUMBER: TSRI 465.0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 784-2937
; TELEFAX: (619) 784-9399
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 571 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-08-779-764A-13
```

```
; APPLICATION NUMBER: US/08/779,764A
; FILING DATE: 16-DEC-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitting, Thomas
; REGISTRATION NUMBER: 34,163
; REFERENCE/DOCKET NUMBER: TSRI 465.0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 784-2937
; TELEFAX: (619) 784-9399
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 571 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-08-779-764A-13

Query Match          63.2%; Score 15.8; DB 3; Length 571;
Best Local Similarity 73.9%; Pred. No. 7.9e+02;
Matches 17; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAGCACACGAGCGCCTC 23
   ||||| ||||| ||||| ||||| |||||
Db 157 AAAAGACACACGCGGTGATCTC 179

RESULT 43
US-09-563-456-13
; Sequence 13, Application US/09563456
; Patent No. 6653464
; GENERAL INFORMATION:
; APPLICANT: de la Torre, Juan C.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR SCREENING
; TITLE OF INVENTION: OF HUMAN BORNA DISEASE VIRUS
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: THE SCRIPPS RESEARCH INSTITUTE
; STREET: 10550 No. 6653464th Torrey Pines Road, TPC-8
; CITY: La Jolla
; STATE: California
; COUNTRY: United States
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/779,764
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitting, Thomas
; REGISTRATION NUMBER: 34,163
; REFERENCE/DOCKET NUMBER: TSRI 465.0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 784-2937
; TELEFAX: (619) 784-9399
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 571 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
```



```
Matches 17; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAACGACAGCGAGCCCTC 23
    ||||| :||:||||| |||
Db 415 AAAAGACRACRCAGCGTGATCTC 437

RESULT 47
US-09-563-456-1
; Sequence 1, Application US/09563456
; Patent No. 6653464
; GENERAL INFORMATION:
; APPLICANT: de la Torre, Juan C.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR SCREENING
; TITLE OF INVENTION: OF HUMAN BORNA DISEASE VIRUS
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: THE SCRIPPS RESEARCH INSTITUTE
; STREET: 10550 No. 6653464th Torrey Pines Road, TPC-8
; CITY: La Jolla
; STATE: California
; COUNTRY: United States
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/563,456
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/779,764
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitting, Thomas
; REGISTRATION NUMBER: 34,163
; REFERENCE/DOCKET NUMBER: TSRI 465.0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 784-2937
; TELEFAX: (619) 784-9399
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8910 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-09-563-456-1

Query Match 63.2%; Score 15.8; DB 3; Length 8910;
Best Local Similarity 73.9%; Pred. No. 9.3e+02;
Matches 17; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAACGACAGCGAGCCCTC 23
    ||||| :||:||||| |||
Db 415 AAAAGACRACRCAGCGTGATCTC 437

RESULT 48
US-09-949-016-17073/c
; Sequence 17073, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17073
; LENGTH: 72662
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(72662)
; OTHER INFORMATION: n = A,T,C or G
; US-09-949-016-17073

Query Match 63.2%; Score 15.8; DB 3; Length 72662;
Best Local Similarity 89.5%; Pred. No. 1.1e+03;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 ACGACACAGCGAGCCCTCG 24
    ||||| ||||| ||||| |||||
Db 16179 ACAACACAGCGAGACCTCG 16161

RESULT 49
US-09-949-016-13734
; Sequence 13734, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13734
; LENGTH: 131254
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(131254)
; OTHER INFORMATION: n = A,T,C or G
; US-09-949-016-13734

Query Match 63.2%; Score 15.8; DB 3; Length 131254;
Best Local Similarity 89.5%; Pred. No. 1.1e+03;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 ACGACACAGCGAGCCCTCG 24
    ||||| ||||| ||||| |||||
Db 67235 ACGACAAAGCGAGCCTTCG 67253

RESULT 50
US-09-949-016-14896
; Sequence 14896, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
```

FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 14896
LENGTH: 135030
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(135030)
OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14896

Query Match 63.2%; Score 15.8; DB 3; Length 135030;
Best Local Similarity 89.5%; Pred. No. 1.1e+03;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 5 AACGACACGCGCCCTC 23
Db 36869 AACAAACACGCGACCTC 36887

Search completed: February 3, 2006, 16:32:18
Job time : 86.6667 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 3, 2006, 14:37:28 ; Search time 365.556 Seconds
(without alignments)
565.535 Million cell updates/sec

Title: US-10-719-900-15

Perfect score: 25

Sequence: 1 aaaaacagacacagcgcctcgg 25

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database : Published Applications NA Main:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	25	100.0	25	8	US-10-719-900-15
2	23.4	93.6	25	8	Sequence 15, Appl
3	19.8	79.2	770	8	Sequence 16, Appl
4	19.8	79.2	770	8	Sequence 13473, A
5	19.8	79.2	770	8	Sequence 13474, A
6	19.8	79.2	770	9	Sequence 13473, A
7	19.8	79.2	770	9	Sequence 13474, A
8	18.6	74.4	352	8	Sequence 24, Appl
9	18.6	74.4	378	7	Sequence 138645, A
10	18.2	72.8	519	5	Sequence 28486, A
11	18.2	72.8	519	5	Sequence 46652, A
12	17.8	71.2	1368	3	Sequence 46652, A
13	17.8	71.2	115780	7	Sequence 893, Appl
14	17.6	70.4	215	7	Sequence 96, Appl
15	17.6	70.4	559	3	Sequence 95630, A
16	17.6	70.4	597	3	Sequence 67625, Ap
17	17.6	70.4	753	8	Sequence 67625, Ap
18	17.6	70.4	753	8	Sequence 26299, A
19	17.6	70.4	753	8	Sequence 26300, A
20	17.6	70.4	753	9	Sequence 26300, A
21	17.6	70.4	1017	8	Sequence 26300, A
22	17.6	70.4	1017	8	Sequence 28097, A
23	17.6	70.4	1017	9	Sequence 28098, A
					Sequence 28097, A

24	17.6	70.4	1017	9	US-10-363-483A-28098	Sequence 28098, A
25	17.6	70.4	1445	8	US-10-425-115-101462	Sequence 101462, A
26	17.6	70.4	6073	8	US-10-357-930-22662	Sequence 22662, A
27	17.6	70.4	6073	8	US-10-357-930-28508	Sequence 28508, A
28	17.2	68.8	597	4	US-09-925-065A-293623	Sequence 293623, A
29	17.2	68.8	1092	4	US-09-925-065A-293624	Sequence 293624, A
30	17.2	68.8	1797	6	US-10-369-493-43475	Sequence 43475, A
31	17.2	68.8	6240	7	US-10-221-596B-22	Sequence 22, Appl
32	17.2	68.8	29686	8	US-10-741-600-19573	Sequence 19573, A
33	17.2	68.8	447	7	US-10-424-599-19441	Sequence 19441, A
34	17.6	68.0	493	7	US-10-425-115-10108	Sequence 10108, A
35	17.6	68.0	602	7	US-10-663-561-333	Sequence 333, App
36	17.6	68.0	615	5	US-10-027-632-189552	Sequence 189552, A
37	17.6	68.0	615	5	US-10-027-632-189553	Sequence 189553, A
38	17.6	68.0	615	5	US-10-027-632-189552	Sequence 189552, A
39	17.6	68.0	615	6	US-10-027-632-189553	Sequence 189553, A
40	17.6	68.0	619	4	US-09-925-065A-803090	Sequence 803090, A
41	17.6	68.0	646	5	US-10-027-632-43803	Sequence 43803, A
42	17.6	68.0	646	5	US-10-027-632-63884	Sequence 63884, A
43	17.6	68.0	646	5	US-10-027-632-179577	Sequence 179577, A
44	17.6	68.0	646	5	US-10-027-632-310475	Sequence 310475, A
45	17.6	68.0	646	6	US-10-027-632-43803	Sequence 43803, A
46	17.6	68.0	646	6	US-10-027-632-63884	Sequence 63884, A
47	17.6	68.0	646	6	US-10-027-632-179577	Sequence 179577, A
48	17.6	68.0	646	6	US-10-027-632-310475	Sequence 310475, A
49	17.6	68.0	6092	9	US-10-450-763-825	Sequence 825, App
50	17.6	68.0	17431	3	US-09-764-855-247	Sequence 247, App
51	17.6	68.0	17431	5	US-10-072-349-247	Sequence 247, App
52	17.6	68.0	18486	8	US-10-719-993-7025	Sequence 7025, App
53	17.6	68.0	29177	8	US-10-719-993-7028	Sequence 7028, App
54	17.6	68.0	38794	9	US-10-965-017-7	Sequence 7, Appl
55	17.6	68.0	39669	9	US-10-965-017-1	Sequence 1, Appl
56	17.6	68.0	44457	8	US-10-719-993-6816	Sequence 6816, App
57	17.6	68.0	62278	7	US-10-322-281-849	Sequence 849, App
58	17.6	68.0	115284	5	US-10-087-192-673	Sequence 673, App
59	17.6	68.0	163701	7	US-10-322-281-439	Sequence 439, App
60	17.6	68.0	198161	8	US-10-775-169-52	Sequence 52, Appl
61	17.6	68.0	198161	8	US-10-723-860-165	Sequence 165, App
62	16.8	67.2	176	7	US-10-767-701-21097	Sequence 21097, A
63	16.8	67.2	272	7	US-10-424-599-120035	Sequence 120035, A
64	16.8	67.2	600	4	US-09-925-065A-366050	Sequence 366050, A
65	16.8	67.2	603	4	US-09-925-065A-599525	Sequence 599525, A
66	16.8	67.2	2217	8	US-10-723-860-6042	Sequence 6042, App
67	16.6	66.4	354	8	US-10-357-930-14297	Sequence 14297, A
68	16.6	66.4	399	8	US-10-357-930-5128	Sequence 5128, App
69	16.6	66.4	401	8	US-10-723-860-3299	Sequence 3299, App
70	16.6	66.4	426	8	US-10-425-115-156587	Sequence 156587, A
71	16.6	66.4	453	7	US-10-424-599-82579	Sequence 82579, A
72	16.6	66.4	459	8	US-10-357-930-35411	Sequence 35411, A
73	16.6	66.4	459	8	US-10-767-701-44243	Sequence 44243, A
74	16.6	66.4	505	7	US-10-767-701-22360	Sequence 22360, A
75	16.6	66.4	547	4	US-09-925-065A-155959	Sequence 155959, A
76	16.6	66.4	547	4	US-09-925-065A-155960	Sequence 155960, A
77	16.6	66.4	582	5	US-10-027-632-194803	Sequence 194803, A
78	16.6	66.4	582	6	US-10-027-632-194803	Sequence 194803, A
79	16.6	66.4	593	4	US-09-925-065A-231068	Sequence 231068, A
80	16.6	66.4	604	4	US-09-925-065A-859836	Sequence 859836, A
81	16.6	66.4	604	4	US-09-925-065A-859837	Sequence 859837, A
82	16.6	66.4	616	4	US-09-925-065A-35438	Sequence 35438, A
83	16.6	66.4	633	4	US-09-925-065A-936672	Sequence 936672, A
84	16.6	66.4	645	4	US-09-925-065A-885836	Sequence 885836, A
85	16.6	66.4	645	4	US-09-925-065A-911136	Sequence 911136, A
86	16.6	66.4	654	4	US-09-925-065A-885837	Sequence 885837, A
87	16.6	66.4	783	8	US-10-363-45A-31779	Sequence 31779, A
88	16.6	66.4	783	8	US-10-363-45A-31780	Sequence 31780, A
89	16.6	66.4	783	9	US-10-363-483A-31779	Sequence 31779, A
90	16.6	66.4	812	5	US-10-363-483A-31780	Sequence 31780, A
91	16.6	66.4	812	5	US-10-027-632-143265	Sequence 143265, A
92	16.6	66.4	812	5	US-10-027-632-143266	Sequence 143266, A
93	16.6	66.4	812	6	US-10-027-632-143265	Sequence 143265, A
94	16.6	66.4	812	6	US-10-027-632-143266	Sequence 143266, A
95	16.6	66.4	922	5	US-10-027-632-164136	Sequence 164136, A
96	16.6	66.4	922	6	US-10-027-632-164136	Sequence 164136, A

Db 409 AAAAAACGACACCGCGACCCCTC 387

RESULT 4

US-10-363-345A-13474
; Sequence 13474, Application US/10363345A
; Publication No. US20040234960A1

; GENERAL INFORMATION:

; APPLICANT: Alexander Olek

; APPLICANT: Christian Piepenbrock

; APPLICANT: Kurt Berlin

; TITLE OF INVENTION: Method for determining the degree of methylation of defined
; TITLE OF INVENTION: cytosines in genomic DNA in the sequence context of 5'-CpG-3

; FILE REFERENCE: E01/1227

; CURRENT APPLICATION NUMBER: US/10/363.345A

; CURRENT FILING DATE: 2003-03-03

; NUMBER OF SEQ ID NOS: 40712

; SEQ ID NO 13474

; LENGTH: 770

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)

; OTHER INFORMATION: CpG-island No: 13474

US-10-363-345A-13474

Query Match 79.2%; Score 19.8; DB 8; Length 770;

Best Local Similarity 91.3%; Pred. No. 29;

Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AAAAAACGACACCGCGACCCCTC 23

Db 362 AAAAAACGACACCGCGACCCCTC 384

RESULT 5

US-10-363-483A-13473/c

; Sequence 13473, Application US/10363483A

; Publication No. US20050064401A1

; GENERAL INFORMATION:

; APPLICANT: Alexander Olek

; APPLICANT: Christian Piepenbrock

; APPLICANT: Kurt Berlin

; TITLE OF INVENTION: Diagnosis of illnesses or predisposition to certain

; TITLE OF INVENTION: illnesses

; FILE REFERENCE: 82011

; CURRENT APPLICATION NUMBER: US/10/363.483A

; CURRENT FILING DATE: 2003-03-03

; NUMBER OF SEQ ID NOS: 40712

; SEQ ID NO 13473

; LENGTH: 770

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)

; OTHER INFORMATION: CpG-island No: 13473

US-10-363-483A-13473

Query Match

Best Local Similarity 79.2%; Score 19.8; DB 9; Length 770;

Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AAAAAACGACACCGCGACCCCTC 23

Db 409 AAAAAACGACACCGCGACCCCTC 387

RESULT 6

US-10-363-483A-13474

; Sequence 13474, Application US/10363483A

; Publication No. US20050064401A1

; GENERAL INFORMATION:

; APPLICANT: Alexander Olek

; APPLICANT: Christian Piepenbrock

; APPLICANT: Kurt Berlin

; TITLE OF INVENTION: Diagnosis of illnesses or predisposition to certain

; TITLE OF INVENTION: illnesses

; FILE REFERENCE: 82011

; CURRENT APPLICATION NUMBER: US/10/363.483A

; CURRENT FILING DATE: 2003-03-03

; NUMBER OF SEQ ID NOS: 40712

; SEQ ID NO 13474

; LENGTH: 770

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)

; OTHER INFORMATION: CpG-island No: 13474

US-10-363-483A-13474

Query Match 79.2%; Score 19.8; DB 9; Length 770;

Best Local Similarity 91.3%; Pred. No. 29;

Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AAAAAACGACACCGCGACCCCTC 23

Db 362 AAAAAACGACACCGCGACCCCTC 384

RESULT 7

US-09-910-943-24

; Sequence 24, Application US/09910943

; Patent No. US20020081610A1

; GENERAL INFORMATION:

; APPLICANT: Hemmati-Brivanlou, Ali

; APPLICANT: Altman, Curtis

; TITLE OF INVENTION: Assays and Materials for Embryonic Gene Expression

; FILE REFERENCE: 7529/1G148US1

; CURRENT APPLICATION NUMBER: US/09/910.943

; CURRENT FILING DATE: 2001-07-23

; NUMBER OF SEQ ID NOS: 742

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 24

; LENGTH: 763

; TYPE: DNA

; ORGANISM: Xenopus laevis

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (1)-(763)

; OTHER INFORMATION: n may be a or g or c or t/u

US-09-910-943-24

Query Match

Best Local Similarity 75.2%; Score 18.8; DB 3; Length 763;

Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 AAACGACACGACCGACCCCTCGG 25

Db 549 AAACGACACGACGAGAACCCCTCGG 570

RESULT 8

US-10-425-115-138645

; Sequence 138645, Application US/10425115

; Publication No. US20040214272A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants

; FILE REFERENCE: 38-21(53222)B

; CURRENT APPLICATION NUMBER: US/10/425.115

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 369326

US-09-974-300-893/c
; Sequence 893, Application US/09974300
; Patent No. US20020146721A1
; GENERAL INFORMATION:
; APPLICANT: Berka, Randy M.
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; FILE REFERENCE: 10085, 500-US
; CURRENT APPLICATION NUMBER: US/09/974,300
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 09/680,598
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/279,526
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 8481
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 893
; LENGTH: 1368
; TYPE: DNA
; ORGANISM: Bacillus licheniformis
US-09-974-300-893

Query Match 71.2%; Score 17.8; DB 3; Length 1368;
Best Local Similarity 90.5%; Pred. No. 2.6e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AAAAAGCACACGAGCCCT 22
|||||
DB 1250 AAAAAGCACAAAGGAGCCCT 1230

RESULT 13
US-10-367-094-96/c
; Sequence 96, Application US/10367094
; Publication No. US20040170982A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc Malandro
; TITLE OF INVENTION: Novel Therapeutic Targets in Cancer
; FILE REFERENCE: 529452001500
; CURRENT APPLICATION NUMBER: US/10/367,094
; CURRENT FILING DATE: 2003-02-14
; NUMBER OF SEQ ID NOS: 203
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 96
; LENGTH: 115780
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(115780)
; OTHER INFORMATION: n = A,T,C or G
US-10-367-094-96

Query Match 71.2%; Score 17.8; DB 7; Length 115780;
Best Local Similarity 90.5%; Pred. No. 3e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 AACGACACACGAGCCCTCGG 25
|||||
DB 111059 AAGGACAAAGGAGCCCTCGG 111039

RESULT 14
US-10-424-599-95630/c
; Sequence 95630, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei

; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 95630
; LENGTH: 215
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_57365C.1
US-10-424-599-95630

Query Match 70.4%; Score 17.6; DB 7; Length 215;
Best Local Similarity 83.3%; Pred. No. 3e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAAGCACACGAGCCCTCG 24
|||||
DB 32 AAAAAAGCACCAAGTGCCTCG 9

RESULT 15
US-09-974-300-6825
; Sequence 6825, Application US/09974300
; Patent No. US20020146721A1
; GENERAL INFORMATION:
; APPLICANT: Berka, Randy M.
; APPLICANT: Clausen, Ib Groth
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; FILE REFERENCE: 10085, 500-US
; CURRENT APPLICATION NUMBER: US/09/974,300
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 09/680,598
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/279,526
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 8481
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6825
; LENGTH: 559
; TYPE: DNA
; ORGANISM: Bacillus clausii
US-09-974-300-6825

Query Match 70.4%; Score 17.6; DB 3; Length 559;
Best Local Similarity 83.3%; Pred. No. 3.1e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAAGCACACGAGCCCTCG 24
|||||
DB 203 AAAAAAGCACATCGAAGCTTCG 226

RESULT 16
US-09-974-300-6762
; Sequence 6762, Application US/09974300
; Patent No. US20020146721A1
; GENERAL INFORMATION:
; APPLICANT: Berka, Randy M.
; APPLICANT: Clausen, Ib Groth
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; FILE REFERENCE: 10085, 500-US
; CURRENT APPLICATION NUMBER: US/09/974,300
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 09/680,598
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/279,526
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 8481

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 6762

; LENGTH: 597

; TYPE: DNA

; ORGANISM: Bacillus clausii

US-09-974-300-6762

Query Match 70.4%; Score 17.6; DB 3; Length 597;

Best Local Similarity 83.3%; Pred. No. 3.1e+02;

Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAACGACACGAGCCCTCG 24

Db 243 AAAAAACGACATCGAAGCTTCG 266

RESULT 17

US-10-363-345A-26299/c

; Sequence 26299, Application US/10363345A

; Publication No. US20040234960A1

; GENERAL INFORMATION:

; APPLICANT: Alexander Olek

; APPLICANT: Christian Piepenbrock

; APPLICANT: Kurt Berlin

; TITLE OF INVENTION: Method for determining the degree of methylation of defined

; FILE REFERENCE: E01/1227

; CURRENT APPLICATION NUMBER: US/10/363,345A

; CURRENT FILING DATE: 2003-03-03

; NUMBER OF SEQ ID NOS: 40712

; SEQ ID NO 26299

; LENGTH: 753

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)

; OTHER INFORMATION: CpG-island No: 26299

US-10-363-345A-26299

Query Match 70.4%; Score 17.6; DB 8; Length 753;

Best Local Similarity 83.3%; Pred. No. 3.1e+02;

Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAACGACACGAGCCCTCG 24

Db 468 AAAAAACGACCCGCGACCCCGCG 445

RESULT 18

US-10-363-345A-26300

; Sequence 26300, Application US/10363345A

; Publication No. US20040234960A1

; GENERAL INFORMATION:

; APPLICANT: Alexander Olek

; APPLICANT: Christian Piepenbrock

; APPLICANT: Kurt Berlin

; TITLE OF INVENTION: Method for determining the degree of methylation of defined

; FILE REFERENCE: E01/1227

; CURRENT APPLICATION NUMBER: US/10/363,345A

; CURRENT FILING DATE: 2003-03-03

; NUMBER OF SEQ ID NOS: 40712

; SEQ ID NO 26300

; LENGTH: 753

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)

; OTHER INFORMATION: CpG-island No: 26300

US-10-363-345A-26300

Query Match 70.4%; Score 17.6; DB 8; Length 753;

Best Local Similarity 83.3%; Pred. No. 3.1e+02;

Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAACGACACGAGCCCTCG 24

Db 286 AAAAAACGACCCGCGACCCCGCG 309

RESULT 19

US-10-363-483A-26299/c

; Sequence 26299, Application US/10363483A

; Publication No. US20050064401A1

; GENERAL INFORMATION:

; APPLICANT: Alexander Olek

; APPLICANT: Christian Piepenbrock

; APPLICANT: Kurt Berlin

; TITLE OF INVENTION: Diagnosis of illnesses or predisposition to certain

; FILE REFERENCE: 82011

; CURRENT APPLICATION NUMBER: US/10/363,483A

; CURRENT FILING DATE: 2003-03-03

; NUMBER OF SEQ ID NOS: 40712

; SEQ ID NO 26299

; LENGTH: 753

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)

; OTHER INFORMATION: CpG-island No: 26299

US-10-363-483A-26299

Query Match 70.4%; Score 17.6; DB 9; Length 753;

Best Local Similarity 83.3%; Pred. No. 3.1e+02;

Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAACGACACGAGCCCTCG 24

Db 468 AAAAAACGACCCGCGACCCCGCG 445

RESULT 20

US-10-363-483A-26300

; Sequence 26300, Application US/10363483A

; Publication No. US20050064401A1

; GENERAL INFORMATION:

; APPLICANT: Alexander Olek

; APPLICANT: Christian Piepenbrock

; APPLICANT: Kurt Berlin

; TITLE OF INVENTION: Diagnosis of illnesses or predisposition to certain

; FILE REFERENCE: 82011

; CURRENT APPLICATION NUMBER: US/10/363,483A

; CURRENT FILING DATE: 2003-03-03

; NUMBER OF SEQ ID NOS: 40712

; SEQ ID NO 26300

; LENGTH: 753

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)

; OTHER INFORMATION: CpG-island No: 26300

US-10-363-483A-26300

Query Match 70.4%; Score 17.6; DB 9; Length 753;

Best Local Similarity 83.3%; Pred. No. 3.1e+02;

Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAACGACACGAGCCCTCG 24

Db 286 AAAAAACGACCCGCGACCCCGCG 309

RESULT 21

US-10-363-345A-28097/c

; Sequence 28097, Application US/10363345A
; Publication No. US20040234960A1
; GENERAL INFORMATION:
; APPLICANT: Alexander Olek
; APPLICANT: Christian Piepenbrock
; TITLE OF INVENTION: Method for determining the degree of methylation of defined
; TITLE OF INVENTION: cytosines in genomic DNA in the sequence context of 5'-CpG-3
; FILE REFERENCE: E01/1227
; CURRENT APPLICATION NUMBER: US/10/363,345A
; CURRENT FILING DATE: 2003-03-03
; NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 28097
; LENGTH: 1017
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; OTHER INFORMATION: CpG-island No: 28097
US-10-363-345A-28097

Query Match 70.4%; Score 17.6; DB 8; Length 1017;
Best Local Similarity 83.3%; Pred. No. 3.1e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAACGACACCGAGCCCTCG 24
|||||
Db 134 AAAAAACGACACCGGTACCCACG 111

RESULT 22

US-10-363-345A-28098
; Sequence 28098, Application US/10363345A
; Publication No. US20040234960A1
; GENERAL INFORMATION:
; APPLICANT: Alexander Olek
; APPLICANT: Christian Piepenbrock
; TITLE OF INVENTION: Method for determining the degree of methylation of defined
; TITLE OF INVENTION: cytosines in genomic DNA in the sequence context of 5'-CpG-3
; FILE REFERENCE: E01/1227
; CURRENT APPLICATION NUMBER: US/10/363,345A
; CURRENT FILING DATE: 2003-03-03
; NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 28098
; LENGTH: 1017
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; OTHER INFORMATION: CpG-island No: 28098
US-10-363-345A-28098

Query Match 70.4%; Score 17.6; DB 8; Length 1017;
Best Local Similarity 83.3%; Pred. No. 3.1e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAACGACACCGAGCCCTCG 24
|||||
Db 884 AAAAAACGACACCGGTACCCACG 907

RESULT 23

US-10-363-483A-28097/c
; Sequence 28097, Application US/10363483A
; Publication No. US20050064401A1
; GENERAL INFORMATION:
; APPLICANT: Alexander Olek
; APPLICANT: Christian Piepenbrock
; TITLE OF INVENTION: Diagnosis of illnesses or predisposition to certain
; TITLE OF INVENTION: illnesses
; FILE REFERENCE: 82011

; CURRENT APPLICATION NUMBER: US/10/363,483A
; CURRENT FILING DATE: 2003-03-03
; NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 28097
; LENGTH: 1017
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; OTHER INFORMATION: CpG-island No: 28097
US-10-363-483A-28097

Query Match 70.4%; Score 17.6; DB 9; Length 1017;
Best Local Similarity 83.3%; Pred. No. 3.1e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAACGACACCGAGCCCTCG 24
|||||
Db 134 AAAAAACGACACCGGTACCCACG 111

RESULT 24

US-10-363-483A-28098
; Sequence 28098, Application US/10363483A
; Publication No. US20050064401A1
; GENERAL INFORMATION:
; APPLICANT: Alexander Olek
; APPLICANT: Christian Piepenbrock
; APPLICANT: Kurt Berlin
; TITLE OF INVENTION: Diagnosis of illnesses or predisposition to certain
; TITLE OF INVENTION: illnesses
; FILE REFERENCE: 82011
; CURRENT APPLICATION NUMBER: US/10/363,483A
; CURRENT FILING DATE: 2003-03-03
; NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 28098
; LENGTH: 1017
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; OTHER INFORMATION: CpG-island No: 28098
US-10-363-483A-28098

Query Match 70.4%; Score 17.6; DB 9; Length 1017;
Best Local Similarity 83.3%; Pred. No. 3.1e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAACGACACCGAGCCCTCG 24
|||||
Db 884 AAAAAACGACACCGGTACCCACG 907

RESULT 25

US-10-425-115-101462
; Sequence 101462, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 101462
; LENGTH: 1445
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:

```
; OTHER INFORMATION: Clone ID: MRT4577_24045C.1
US-10-425-115-101462

Query Match          70.4%; Score 17.6; DB 8; Length 1445;
Best Local Similarity 83.3%; Pred. No. 3.2e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AAAAAGCAGACAGCGAGCCCTCG 25
   ||| ||||| ||||| ||||| |||||
Db 48 AAAAGCAGACAGCGAGCGGCGG 71

RESULT 26
US-10-357-930-22662
; Sequence 22662, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22662
; LENGTH: 6073
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 6067, 6068, 6069, 6070, 6071, 6072, 6073
; OTHER INFORMATION: n = A,T,C or G
US-10-357-930-22662

Query Match          70.4%; Score 17.6; DB 8; Length 6073;
Best Local Similarity 83.3%; Pred. No. 3.3e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAGCAGACAGCGAGCCCTCG 24
   ||||| ||| ||| ||||| |||||
Db 12 AAAAGAGACTCGCGAGCCCTCG 35

RESULT 27
US-10-357-930-28508
; Sequence 28508, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
US-10-357-930-28508

Query Match          70.4%; Score 17.6; DB 8; Length 6073;
Best Local Similarity 83.3%; Pred. No. 3.3e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAGCAGACAGCGAGCCCTCG 24
   ||||| ||| ||| ||||| |||||
Db 12 AAAAGAGACTCGCGAGCCCTCG 35

RESULT 28
US-09-925-065A-293623/c
; Sequence 293623, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 10827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR APPLICATION NUMBER: 2001-08-08
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 293623
; LENGTH: 597
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(597)
; OTHER INFORMATION: n = A,T,C or G
US-09-925-065A-293623

Query Match          68.8%; Score 17.2; DB 4; Length 597;
Best Local Similarity 86.4%; Pred. No. 4.7e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

QY 1 AAAAAACGACACAGCGGCCT 22
|||||
Db 580 AAAAAACACAAAGCGGCACT 559

RESULT 29
US-09-925-065A-293624/c
; Sequence 293624, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108927.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 293624
; LENGTH: 1092
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(1092)
; OTHER INFORMATION: n = A,T,C or G
US-09-925-065A-293624

Query Match 68.8%; Score 17.2; DB 4; Length 1092;
Best Local Similarity 86.4%; Pred. No. 4.8e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAACGACACAGCGGCCT 22
|||||
Db 580 AAAAAACACAAAGCGGCACT 559

RESULT 30
US-10-369-493-43475
; Sequence 43475, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 43475
; LENGTH: 1797
; TYPE: DNA
; ORGANISM: No. US20030233675Altoc punctiforme
US-10-369-493-43475

Query Match 68.8%; Score 17.2; DB 6; Length 1797;

Best Local Similarity 86.4%; Pred. No. 4.9e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 AAAAAACGACACAGCGGCCT 22
|||||
Db 336 AAAAAACGAAACAGCGATCGCT 357

RESULT 31
US-10-221-596B-22/c
; Sequence 22, Application US/10221596B
; Publication No. US20040086855A1
; GENERAL INFORMATION:
; APPLICANT: Iwamoto, Masao
; APPLICANT: Higo, Kenichi
; TITLE OF INVENTION: Method for Screening Genes Expressing at Desired Sites
; FILE REFERENCE: 59150-8020
; CURRENT APPLICATION NUMBER: US/10/221,596B
; CURRENT FILING DATE: 2003-10-24
; PRIOR APPLICATION NUMBER: PCT/JP01/10195
; PRIOR FILING DATE: 2001-11-21
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 22
; LENGTH: 6240
; TYPE: DNA
; ORGANISM: Oryza sativa
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: AF000559
; DATABASE ENTRY DATE: OCT-1999
; RELEVANT RESIDUES: 7921 TO 14161
US-10-221-596B-22

Query Match 68.8%; Score 17.2; DB 7; Length 6240;
Best Local Similarity 86.4%; Pred. No. 5.1e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAACGACACAGCGGCCT 22
|||||
Db 1786 AAAAAATGACACAGCAAGCCAT 1765

RESULT 32
US-10-741-600-17573
; Sequence 17573, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17573
; LENGTH: 29686
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(29686)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1
US-10-741-600-17573

Query Match 68.8%; Score 17.2; DB 8; Length 29686;
Best Local Similarity 86.4%; Pred. No. 5.4e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAACGACACAGCGGCCT 22
|||||
Db 26855 AAAAAATGACAAAGCTAGCCCT 26876

RESULT 33
 US-10-424-599-19441
 ; Sequence 19441, Application US/10424599
 ; Publication No. US20040031072A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa Thomas J
 ; APPLICANT: Kovalic David K
 ; APPLICANT: Zhou Yihua
 ; APPLICANT: Cao Yongwei
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53223)B
 ; CURRENT APPLICATION NUMBER: US/10/424,599
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 285684
 ; SEQ ID NO 19441
 ; LENGTH: 447
 ; TYPE: DNA
 ; ORGANISM: Glycine max
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: PAT_MRT3847_117559C.1
 US-10-424-599-19441

```
Query Match      68.0%; Score 17; DB 7; Length 447;
Best Local Similarity 80.0%; Pred. No. 5.8e+02;
Matches 20: Conservative 0; Mismatches 5; Indels 0; Gaps 0;
```

Qy 1 AAAAAACGACACAGCGAGCCCTCGG 25
d6 45 AAAAAACGACAGGCGAGCCCTGGG 69

```

RESULT 34
US-10-425-115-10108
; Sequence 10108, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 10108
; LENGTH: 493
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(493)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_10921C.1
US-10-425-115-10108

```

Query Match	68.0%;	Score 17;	DB 8;	Length 493;
Best Local Similarity	80.0%;	Pred. No. 5.8e+02;		
Matches 20:	Conservative	0;	Mismatches 5;	Indels 0;
				Gaps 0;

Qy 1 AAAAAACGACACAGCGGCCCTCGG 25
||| ||| ||| ||| ||| ||| |||
pB 261 AAAAGTCGACACAGCGGCTCTCGG 285

RESULT 35
US-10-663-561-333
; Sequence 333, Application US/106633561
; Publication No. US20040166509A1

```

; GENERAL INFORMATION:
; APPLICANT: DENSLOW, NANCY D.
; APPLICANT: LARKIN, PATRICK M.
; APPLICANT: FOLMAR, LEROY C.
; APPLICANT: HEMMER, MICHAEL J.
; TITLE OF INVENTION: DETECTING HORMONALLY ACTIVE COMPOUNDS
; FILE REFERENCE: 5853-238
; CURRENT APPLICATION NUMBER: US/10/663,561
; CURRENT FILING DATE: 2003-09-15
; NUMBER OF SEQ ID NOS: 566
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 333
; LENGTH: 602
; TYPE: DNA
; ORGANISM: Cyprinodon variegatus
; US-10-663-561-333

```

Query Match	68.0%	Score 17;	DB 7;	Length 602;
Best Local Similarity	80.0%	Pred. No. 5.9e+02;		
Matches	20:	Conservative	0:	Mismatches 5;
		Indels	0:	Gaps 0;

QY 1 AAAAAACGACACAGCGAGCCCTCGG 25
||||| ||| ||| ||| ||| ||| ||| |||
Db 547 AAAAAACAACAAGCTTGACCTCGG 571

```

RESULT 36
US-10-027-632-189552/c
; Sequence 189552, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 189552
; LENGTH: 615
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-189552

```

Query Match 68.0%; Score 17; DB 5; Length 615;
Best Local Similarity 80.0%; Pred. No. 5.9e+02;
Matches 20; Conservative 0; Mismatches 5; Indels

Qy 1 AAAAAACGACACAGCGAGCCCTCGG 25
||| ||| ||| ||| ||| ||| ||| |||
Db 569 AGAAAGGACACAGCCAGCGCGCGG 545

RESULT 37
US-10-027-632-189553/c
; Sequence 189553, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:

APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 189553
; LENGTH: 615
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-189553

Query Match 68.0%; Score 17; DB 5; Length 615;
Best Local Similarity 80.0%; Pred. No. 5.9e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 AAAAAACGACACAGCGAGCCCTCGG 25
Db 569 AGAAAAGGACACAGCGAGCCCGCGG 545

RESULT 38
US-10-027-632-189552/c
; Sequence 189552, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 189552
; LENGTH: 615
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-189552

Query Match 68.0%; Score 17; DB 6; Length 615;
Best Local Similarity 80.0%; Pred. No. 5.9e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 AAAAAACGACACAGCGAGCCCTCGG 25
Db 569 AGAAAAGGACACAGCGAGCCCGCGG 545

RESULT 39
US-10-027-632-189553/c
; Sequence 189553, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 189553
; LENGTH: 615
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-189553

Query Match 68.0%; Score 17; DB 6; Length 615;
Best Local Similarity 80.0%; Pred. No. 5.9e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 AAAAAACGACACAGCGAGCCCTCGG 25
Db 569 AGAAAAGGACACAGCGAGCCCGCGG 545

RESULT 40
US-09-925-065A-803090/c
; Sequence 803090, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 803090
; LENGTH: 619

```
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-803090

Query Match      68.0%; Score 17; DB 4; Length 619;
Best Local Similarity 80.0%; Pred. No. 5.9e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 AAAAAACGACACAGCGAGCCCTCGG 25
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 217 AAAAAATAGACACAGACGACCCCTCAG 193

RESULT 41
US-10-027-632-43803
; Sequence 43803, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 43803
; LENGTH: 646
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-43803

Query Match      68.0%; Score 17; DB 5; Length 646;
Best Local Similarity 80.0%; Pred. No. 5.9e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 AAAAAACGACACAGCGAGCCCTCGG 25
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 355 AAAGATCCACACAGTGAGCCCTCTG 379

RESULT 42
US-10-027-632-63884
; Sequence 63884, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 43803
; LENGTH: 646
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-43803

Query Match      68.0%; Score 17; DB 5; Length 646;
Best Local Similarity 80.0%; Pred. No. 5.9e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 AAAAAACGACACAGCGAGCCCTCGG 25
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 355 AAAGATCCACACAGTGAGCCCTCTG 379

RESULT 43
US-10-027-632-179577
; Sequence 179577, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 179577
; LENGTH: 646
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-179577

Query Match      68.0%; Score 17; DB 5; Length 646;
Best Local Similarity 80.0%; Pred. No. 5.9e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 AAAAAACGACACAGCGAGCCCTCGG 25
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 355 AAAGATCCACACAGTGAGCCCTCTG 379

RESULT 44
US-10-027-632-179577
; Sequence 179577, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 179577
; LENGTH: 646
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-179577

Query Match      68.0%; Score 17; DB 5; Length 646;
Best Local Similarity 80.0%; Pred. No. 5.9e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 AAAAAACGACACAGCGAGCCCTCGG 25
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 355 AAAGATCCACACAGTGAGCCCTCTG 379
```



```
RESULT 44
US-10-027-632-310475
; Sequence 310475, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; POLYMORPHISMS IN THE HUMAN GENOME
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 310475
; LENGTH: 646
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(646)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-310475

Query Match          68.0%; Score 17; DB 5; Length 646;
Best Local Similarity 80.0%; Pred. No. 5.9e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY      1 AAAAAACGACACAGCGAGCCCTCGG 25
        ||| ||| ||| ||| ||| ||| |||
Db      355 AAAGATCCACACAGTGAGCCCTCTG 379

RESULT 45
US-10-027-632-43803
; Sequence 43803, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; POLYMORPHISMS IN THE HUMAN GENOME
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 43803
; LENGTH: 646
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(646)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-43803

Query Match          68.0%; Score 17; DB 5; Length 646;
Best Local Similarity 80.0%; Pred. No. 5.9e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY      1 AAAAAACGACACAGCGAGCCCTCGG 25
        ||| ||| ||| ||| ||| ||| |||
Db      355 AAAGATCCACACAGTGAGCCCTCTG 379

RESULT 46
US-10-027-632-63884
; Sequence 63884, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; POLYMORPHISMS IN THE HUMAN GENOME
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 63884
; LENGTH: 646
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(646)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-63884

Query Match          68.0%; Score 17; DB 6; Length 646;
Best Local Similarity 80.0%; Pred. No. 5.9e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY      1 AAAAAACGACACAGCGAGCCCTCGG 25
        ||| ||| ||| ||| ||| ||| |||
Db      355 AAAGATCCACACAGTGAGCCCTCTG 379

RESULT 47
US-10-027-632-179577
; Sequence 179577, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; POLYMORPHISMS IN THE HUMAN GENOME
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
```

;; PRIOR APPLICATION NUMBER: US 60/218,006
;; PRIOR FILING DATE: 2000-07-12
;; PRIOR APPLICATION NUMBER: US 60/198,676
;; PRIOR FILING DATE: 2000-04-20
;; PRIOR APPLICATION NUMBER: US 60/193,483
;; PRIOR FILING DATE: 2000-03-29
;; PRIOR APPLICATION NUMBER: US 60/185,218
;; PRIOR FILING DATE: 2000-02-24
;; PRIOR APPLICATION NUMBER: US 60/167,363
;; PRIOR FILING DATE: 1999-11-23
;; PRIOR APPLICATION NUMBER: US 60/156,358
;; PRIOR FILING DATE: 1999-09-28
;; PRIOR APPLICATION NUMBER: US 60/146,002
;; PRIOR FILING DATE: 1999-08-09
;; NUMBER OF SEQ ID NOS: 325720
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 179577
;; LENGTH: 646
;; TYPE: DNA
;; ORGANISM: Human
;; FEATURE:
;; NAME/KEY: misc_feature
;; LOCATION: (1)...(646)
;; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-179577

Query Match 68.0%; Score 17; DB 6; Length 646;
Best Local Similarity 80.0%; Pred. No. 5.9e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAAACGACACGAGCCCTCGG 25
DB 355 AAAGATCCACAGTGAGCCCTCTG 379

RESULT 48
US-10-027-632-310475
;; Sequence 310475, Application US/10027632
;; Publication No. US20030204075A9
;; GENERAL INFORMATION:
;; APPLICANT: Wang, David G.
;; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;; POLYMORPHISMS IN THE HUMAN GENOME
;; FILE REFERENCE: 108827.129
;; CURRENT APPLICATION NUMBER: US/10/027,632
;; CURRENT FILING DATE: 2002-04-30
;; PRIOR APPLICATION NUMBER: US 60/218,006
;; PRIOR FILING DATE: 2000-07-12
;; PRIOR APPLICATION NUMBER: US 60/198,676
;; PRIOR FILING DATE: 2000-04-20
;; PRIOR APPLICATION NUMBER: US 60/193,483
;; PRIOR FILING DATE: 2000-03-29
;; PRIOR APPLICATION NUMBER: US 60/185,218
;; PRIOR FILING DATE: 2000-02-24
;; PRIOR APPLICATION NUMBER: US 60/167,363
;; PRIOR FILING DATE: 1999-11-23
;; PRIOR APPLICATION NUMBER: US 60/156,358
;; PRIOR FILING DATE: 1999-09-28
;; PRIOR APPLICATION NUMBER: US 60/146,002
;; PRIOR FILING DATE: 1999-08-09
;; NUMBER OF SEQ ID NOS: 325720
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 310475
;; LENGTH: 646
;; TYPE: DNA
;; ORGANISM: Human
;; FEATURE:
;; NAME/KEY: misc_feature
;; LOCATION: (1)...(646)
;; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-310475

Query Match 68.0%; Score 17; DB 6; Length 646;

Best Local Similarity 80.0%; Pred. No. 5.9e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 1 AAAAAACGACACGAGCCCTCGG 25
DB 355 AAAGATCCACAGTGAGCCCTCTG 379

RESULT 49
US-10-450-763-825
;; Sequence 825, Application US/10450763
;; Publication No. US20050196754A1
;; GENERAL INFORMATION:
;; APPLICANT: Hyseq, Inc
;; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
;; FILE REFERENCE: 790CIP3/US
;; CURRENT APPLICATION NUMBER: US/10/450,763
;; CURRENT FILING DATE: 2003-06-11
;; PRIOR APPLICATION NUMBER: PCT/US01/08631
;; PRIOR FILING DATE: 2001-03-30
;; PRIOR APPLICATION NUMBER: 09/540,217
;; PRIOR FILING DATE: 2000-03-31
;; PRIOR APPLICATION NUMBER: 09/649,167
;; PRIOR FILING DATE: 2000-08-23
;; NUMBER OF SEQ ID NOS: 60736
;; SOFTWARE: Custom
;; SEQ ID NO 825
;; LENGTH: 6092
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: SIMILAR
;; LOCATION: (6092)..(6069)
;; OTHER INFORMATION: 100% homologous to Homo sapiens Polypeptide identified by the
;; OTHER INFORMATION: signal sequence trap method, accession number Y02377, Smith-Waterman
;; OTHER INFORMATION: Score=44.
US-10-450-763-825

Query Match 68.0%; Score 17; DB 9; Length 6092;
Best Local Similarity 80.0%; Pred. No. 6.3e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAAACGACACGAGCCCTCGG 25
DB 2200 AAAAAAGAGATAGCACCCCTCAG 2224

RESULT 50
US-09-764-855-247
;; Sequence 247, Application US/09764855
;; Patent No. US20020119919A1
;; GENERAL INFORMATION:
;; APPLICANT: Rosen et al.
;; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;; FILE REFERENCE: P4110
;; CURRENT APPLICATION NUMBER: US/09/764,855
;; CURRENT FILING DATE: 2001-01-17
;; Prior application data removed - consult PALM or file wrapper
;; NUMBER OF SEQ ID NOS: 334
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 247
;; LENGTH: 17431
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-09-764-855-247

Query Match 68.0%; Score 17; DB 3; Length 17431;
Best Local Similarity 80.0%; Pred. No. 6.5e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAAACGACACGAGCCCTCGG 25
DB 7370 AAAAAACAAATGCGAGCCCTCTG 7394

Search completed: February 3, 2006, 15:43:46
Job time : 370.556 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 3, 2006, 11:35:31 ; Search time 343.111 Seconds
(without alignments)
61.064 Million cell updates/sec

Title: US-10-719-900-15

Perfect score: 25

Sequence: 1 aaaaacgacacagcagcctcgg 25

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 6068529 seqs, 41903697 residues

Total number of hits satisfying chosen parameters: 12137058

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database : Published Applications NA New:

- 1: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
- 2: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
- 5: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
- 7: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
- 9: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
- 10: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
- 11: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	17.8	71.2	2347	7	US-10-510-386-23
C 2	17.2	68.8	172649	8	US-11-121-086-36
C 3	17	68.0	860	7	US-10-750-185-51708
C 4	17	68.0	860	7	US-10-750-623-51708
C 5	17	68.0	3555	7	US-10-750-185-31750
C 6	17	68.0	3555	7	US-10-750-623-31750
C 7	17	68.0	198161	7	US-10-775-169-52
C 8	16.6	66.4	1177	7	US-10-750-185-40395
C 9	16.6	66.4	1177	7	US-10-750-623-40395
C 10	16.6	66.4	1524	7	US-10-750-185-44232
C 11	16.6	66.4	1524	7	US-10-750-623-44232
C 12	16.2	64.8	201	8	US-11-124-368A-3989
C 13	16	64.0	201	7	US-10-995-561-72930
C 14	16	64.0	1409	7	US-10-750-185-28365
C 15	16	64.0	1409	7	US-10-750-623-28365
C 16	16	64.0	1432	8	US-11-000-688-716
C 17	16	64.0	1505	7	US-10-750-185-36115
C 18	16	64.0	1505	7	US-10-750-623-36115
C 19	16	64.0	60754	7	US-10-995-561-13440
C 20	16	64.0	96128	7	US-10-995-561-13197
C 21	16	64.0	130733	8	US-11-121-086-19
C 22	16	64.0	162085	8	US-11-121-086-7
C 23	16	64.0	164810	8	US-11-121-086-4
C 24	15.8	63.2	25	8	US-11-121-849-379393
C 25	15.8	63.2	25	8	US-11-136-527-352769
C 26	15.8	63.2	1400	8	US-11-136-527-7464
C 27	15.8	63.2	1713	7	US-10-750-185-54749
C 28	15.8	63.2	1713	7	US-10-750-623-54749
C 29	15.8	63.2	2383	8	US-11-136-521-3368
C 30	15.6	62.4	201	7	US-10-995-561-62101
C 31	15.6	62.4	201	7	US-10-995-561-62151
C 32	15.6	62.4	777	7	US-10-750-185-58178
C 33	15.6	62.4	777	7	US-10-750-623-58178
C 34	15.6	62.4	794	7	US-10-750-185-29381
C 35	15.6	62.4	794	7	US-10-750-623-29381
C 36	15.6	62.4	870	8	US-11-041-776-65
C 37	15.6	62.4	1045	7	US-10-750-185-59482
C 38	15.6	62.4	1045	7	US-10-750-623-59482
C 39	15.6	62.4	1197	8	US-11-136-527-132
C 40	15.6	62.4	1490	7	US-10-750-185-44586
C 41	15.6	62.4	1490	7	US-10-750-623-44586
C 42	15.6	62.4	1700	7	US-10-750-185-52468
C 43	15.6	62.4	1700	7	US-10-750-623-52468
C 44	15.6	62.4	1771	7	US-10-750-185-35903
C 45	15.6	62.4	1771	7	US-10-750-623-35903
C 46	15.6	62.4	54767	7	US-10-995-561-13357
C 47	15.6	62.4	151169	8	US-11-121-086-38
C 48	15.4	61.6	55	6	US-10-517-151-105
C 49	15.4	61.6	121	7	US-10-310-914A-845
C 50	15.4	61.6	201	7	US-10-995-561-41053
C 51	15.4	61.6	201	7	US-10-995-561-41056
C 52	15.4	61.6	711	7	US-10-750-185-43072
C 53	15.4	61.6	711	7	US-10-750-623-43072
C 54	15.4	61.6	1060	7	US-10-750-185-58209
C 55	15.4	61.6	1060	7	US-10-750-623-58209
C 56	15.4	61.6	1083	7	US-10-750-185-27234
C 57	15.4	61.6	1083	7	US-10-750-623-27234
C 58	15.4	61.6	1385	7	US-10-750-185-28881
C 59	15.4	61.6	1385	7	US-10-750-623-28881
C 60	15.4	61.6	1734	7	US-10-750-185-37863
C 61	15.4	61.6	1734	7	US-10-750-623-37863
C 62	15.4	61.6	1930	7	US-10-750-185-47014
C 63	15.4	61.6	1930	7	US-10-750-623-47014
C 64	15.4	61.6	2080	7	US-10-750-185-48927
C 65	15.4	61.6	2080	7	US-10-750-623-48927
C 66	15.4	61.6	2626	7	US-10-750-185-58314
C 67	15.4	61.6	2626	7	US-10-750-623-58314
C 68	15.4	61.6	4589	7	US-10-750-185-62517
C 69	15.4	61.6	4589	7	US-10-750-623-62517
C 70	15.4	61.6	4887	7	US-10-750-185-61589
C 71	15.4	61.6	4887	7	US-10-750-623-61589
C 72	15.4	61.6	15042	7	US-10-995-561-13391
C 73	15.4	61.6	86081	7	US-10-995-561-13246
C 74	15.4	61.6	161984	8	US-11-112-908-57
C 75	15.4	61.6	167891	8	US-11-121-086-14
C 76	15.4	61.6	191091	8	US-11-121-086-60
C 77	15.4	61.6	197781	8	US-11-112-908-34
C 78	15.4	61.6	321019	7	US-10-995-561-13204
C 79	15.2	60.8	25	8	US-11-121-849-199027
C 80	15.2	60.8	455	8	US-11-000-688-1532
C 81	15.2	60.8	3119	8	US-11-000-688-398
C 82	15.2	60.8	5798	8	US-11-136-527-3895
C 83	15.2	60.8	20492	8	US-11-095-668-2
C 84	15.2	60.8	127917	7	US-10-775-169-82
C 85	15.2	60.8	1080000	7	US-10-928-446A-1
C 86	15.2	60.8	1080000	7	US-10-928-446A-181
C 87	15.2	60.8	1080000	7	US-10-928-446A-183
C 88	15.2	60.8	1080000	7	US-10-928-446A-185
C 89	15.2	60.8	1080000	7	US-10-928-446A-187
C 90	15.2	60.8	1080000	7	US-10-928-446A-189
C 91	15.2	60.8	1080000	7	US-10-928-446A-191
C 92	15.2	60.8	1080000	7	US-10-928-446A-193
C 93	15.2	60.8	1080000	7	US-10-928-446A-195
C 94	15.2	60.8	1080000	7	US-10-928-446A-197
C 95	15.2	60.8	1080000	7	US-10-928-446A-199

Sequence 4, Appli
Sequence 379393,
Sequence 352769,
Sequence 7464, Ap
Sequence 54749, A
Sequence 54749, A
Sequence 3368, Ap
Sequence 62101, A
Sequence 62151, A
Sequence 58178, A
Sequence 58178, A
Sequence 29381, A
Sequence 29381, A
Sequence 59482, A
Sequence 59482, A
Sequence 132, App
Sequence 44586, A
Sequence 44586, A
Sequence 52468, A
Sequence 52468, A
Sequence 35903, A
Sequence 35903, A
Sequence 13357, A
Sequence 38, Appl
Sequence 105, App
Sequence 845, App
Sequence 41053, A
Sequence 41056, A
Sequence 43072, A
Sequence 43072, A
Sequence 58209, A
Sequence 58209, A
Sequence 27234, A
Sequence 27234, A
Sequence 28881, A
Sequence 28881, A
Sequence 37863, A
Sequence 37863, A
Sequence 47014, A
Sequence 47014, A
Sequence 48927, A
Sequence 48927, A
Sequence 58314, A
Sequence 58314, A
Sequence 62517, A
Sequence 62517, A
Sequence 62517, A
Sequence 61589, A
Sequence 61589, A
Sequence 13391, A
Sequence 13246, A
Sequence 57, Appl
Sequence 14, Appl
Sequence 60, Appl
Sequence 34, Appl
Sequence 13204, A
Sequence 153027, A
Sequence 153027, A
Sequence 388, App
Sequence 3895, App
Sequence 2, Appl
Sequence 82, Appl
Sequence 1, Appl
Sequence 181, App
Sequence 183, App
Sequence 185, App
Sequence 187, App
Sequence 187, App
Sequence 189, App
Sequence 191, App
Sequence 191, App
Sequence 193, App
Sequence 195, App
Sequence 197, App
Sequence 199, App

```
15.2 60.8 1080000 7 US-10-928-446A-201
c 96 15 60.0 201 7 US-10-995-561-8797
c 97 15 60.0 201 7 US-10-995-561-8803
c 98 15 60.0 201 7 US-10-995-561-8809
c 99 15 60.0 201 7 US-10-995-561-8809
c 100 15 60.0 201 7 US-10-995-561-8858
c 101 15 60.0 201 7 US-10-995-561-8864
c 102 15 60.0 201 7 US-10-995-561-8870
c 103 15 60.0 201 7 US-10-995-561-8919
c 104 15 60.0 201 7 US-10-995-561-8925
c 105 15 60.0 201 7 US-10-995-561-8931
c 106 15 60.0 201 7 US-10-995-561-41058
c 107 15 60.0 201 7 US-10-995-561-48229
c 108 15 60.0 201 7 US-10-995-561-48243
c 109 15 60.0 201 7 US-10-995-561-72625
c 110 15 60.0 201 7 US-10-995-561-72642
c 111 15 60.0 201 7 US-10-995-561-84409
c 112 15 60.0 201 7 US-10-995-561-84596
c 113 15 60.0 201 7 US-10-802-796-211
c 114 15 60.0 360 7 US-11-092-140-108
c 115 15 60.0 575 8 US-11-136-527-6296
c 116 15 60.0 600 8 US-11-136-527-6296
c 117 15 60.0 800 7 US-10-750-185-49375
c 118 15 60.0 800 7 US-10-750-623-49375
c 119 15 60.0 1086 7 US-10-750-185-48479
c 120 15 60.0 1086 7 US-10-750-623-48479
c 121 15 60.0 1285 7 US-10-750-623-50362
c 122 15 60.0 1285 7 US-10-750-623-50362
c 123 15 60.0 1785 7 US-10-750-185-54653
c 124 15 60.0 2098 7 US-10-750-623-54653
c 125 15 60.0 2098 7 US-10-750-185-41812
c 126 15 60.0 2175 7 US-10-750-623-41812
c 127 15 60.0 2175 7 US-10-750-185-57980
c 128 15 60.0 2254 8 US-11-136-527-2200
c 129 15 60.0 3624 7 US-10-750-623-62428
c 130 15 60.0 3624 7 US-10-750-623-62428
c 131 15 60.0 4507 8 US-11-136-527-1960
c 132 15 60.0 5383 8 US-11-136-527-143
c 133 15 60.0 5495 7 US-10-775-169-342
c 134 15 60.0 5982 8 US-11-034-771-1
c 135 15 60.0 6421 7 US-10-485-517-27
c 136 15 60.0 7715 8 US-11-147-109-1
c 137 15 60.0 9616 7 US-10-995-561-309
c 138 15 60.0 9626 7 US-10-995-561-308
c 139 15 60.0 9636 7 US-10-995-561-310
c 140 15 60.0 11050 7 US-10-240-708-85
c 141 15 60.0 15804 7 US-10-995-561-13294
c 142 15 60.0 24446 7 US-10-995-561-13436
c 143 15 60.0 153142 8 US-11-121-086-27
c 144 15 60.0 162085 8 US-11-121-086-7
c 145 15 60.0 163162 8 US-11-121-086-66
c 146 15 60.0 185393 8 US-11-121-086-101
c 147 15 60.0 220895 8 US-10-775-169-88
c 148 15 60.0 1125000 7 US-10-995-561-13286
c 149 14.8 59.2 19 9 US-11-101-244-971748
c 150 14.8 59.2 19 10 US-11-083-784-971748
```

ALIGNMENTS

```
RESULT 1
US-10-510-386-23/c
; Sequence 23, Application US/10510386
; Publication No. US20050244922A1
; GENERAL INFORMATION:
; APPLICANT: Andersen, Jens Tonne
; APPLICANT: Clausen, Ib Groth
; APPLICANT: Jorgensen, Steen Troels
; APPLICANT: Olsen, Peter Bjarke
; APPLICANT: Rasmussen, Michael Dolberg
; TITLE OF INVENTION: Improved Bacillus Host Cell
; FILE REFERENCE: 10294.204-US
; CURRENT APPLICATION NUMBER: US/10/510,386
```

```
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 248
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 23
; LENGTH: 2347
; TYPE: DNA
; ORGANISM: Bacillus licheniformis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (513)..(1856)
US-10-510-386-23

Query Match 71.2%; Score 17.8; DB 7; Length 2347;
Best Local Similarity 90.5%; Pred. No. 31;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AAAAAACGACACAGCGAGCCCT 22
|||||
Db 327 AAAAAGACAAAGGAGCCCT 307
|||||

RESULT 2
US-11-121-086-36
; Sequence 36, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138.6000-00000
; CURRENT APPLICATION NUMBER: US/11/121,086
; CURRENT FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 36
; LENGTH: 172649
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-36

Query Match 68.8%; Score 17.2; DB 8; Length 172649;
Best Local Similarity 86.4%; Pred. No. 1.2e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAACGACACAGCGAGCCCT 22
|||||
Db 144092 AAAAATGACAAAGCTAGCCCT 144113
|||||

RESULT 3
US-10-750-185-51708
; Sequence 51708, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: Denise, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 51708
; LENGTH: 860
```

```
; TYPE: DNA
; ORGANISM: Bovine 19866881570598
US-10-750-185-51708

Query Match      68.0%; Score 17; DB 7; Length 860;
Best Local Similarity 80.0%; Pred. No. 65;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAAACGACACGAGCGCCCTCGG 25
Db 728 AAAAAAGTCACAGCAATCCCTCTG 752

RESULT 4
US-10-750-623-51708
; Sequence 51708, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 51708
; LENGTH: 860
; TYPE: DNA
; ORGANISM: Bovine 19866881570598
US-10-750-623-51708

Query Match      68.0%; Score 17; DB 7; Length 860;
Best Local Similarity 80.0%; Pred. No. 65;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAAACGACACGAGCGCCCTCGG 25
Db 728 AAAAAAGTCACAGCAATCCCTCTG 752

RESULT 5
US-10-750-185-31750/c
; Sequence 31750, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 31750
; LENGTH: 3555
; TYPE: DNA
; ORGANISM: Bovine 19866881076987
US-10-750-185-31750

Query Match      68.0%; Score 17; DB 7; Length 860;
Best Local Similarity 80.0%; Pred. No. 65;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAAACGACACGAGCGCCCTCGG 25
Db 728 AAAAAAGTCACAGCAATCCCTCTG 752

RESULT 6
US-10-750-623-31750/c
; Sequence 31750, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 31750
; LENGTH: 3555
; TYPE: DNA
; ORGANISM: Bovine 19866881076987
US-10-750-623-31750

Query Match      68.0%; Score 17; DB 7; Length 3555;
Best Local Similarity 80.0%; Pred. No. 81;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAAACGACACGAGCGCCCTCGG 25
Db 3525 AAGAAACGCGAGAGCAAGCCCTCGG 3501

RESULT 7
US-10-775-169-52/c
; Sequence 52, Application US/10775169
; Publication No. US20050287532A9
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Burczynski, Michael
; APPLICANT: Twine, Natalie
; APPLICANT: Dörner, Andrew
; APPLICANT: Trepicchio, William
; TITLE OF INVENTION: Method for Monitoring Drug Activities In Vivo
; FILE REFERENCE: AM101080 (031896-013000)
; CURRENT APPLICATION NUMBER: US/10775,169
; CURRENT FILING DATE: 2004-02-11
; NUMBER OF SEQ ID NOS: 5278
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 52
; LENGTH: 198161
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-775-169-52

Query Match      68.0%; Score 17; DB 7; Length 198161;
Best Local Similarity 80.0%; Pred. No. 1.5e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAAACGACACGAGCGCCCTCGG 25
Db 198161 AAAAAACGACACGAGCGCCCTCGG 198161
```



```
; Publication No. US20050287559A1
; GENERAL INFORMATION:
; APPLICANT: Michele Cargill
; APPLICANT: James J. Devlin
; APPLICANT: May Luke
; TITLE OF INVENTION: Genetic Polymorphisms Associated with
; TITLE OF INVENTION: Vascular Diseases, Methods of Detection and Uses Thereof
; FILE REFERENCE: CL001524
; CURRENT APPLICATION NUMBER: US/11/124,368A
; CURRENT FILING DATE: 2005-05-09
; PRIOR APPLICATION NUMBER: US 60/568,845
; PRIOR FILING DATE: 2004-05-07
; PRIOR APPLICATION NUMBER: US 60/625,936
; PRIOR FILING DATE: 2004-11-09
; NUMBER OF SEQ ID NOS: 21112
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3989
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-124-368A-3989

Query Match      64.8%; Score 16.2; DB 8; Length 201;
Best Local Similarity 85.7%; Pred. No. 1.2e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy  2 AAAAAGCAGACGAGCCCT 22
Db  67 AAAAACAACCCAGAGCCCT 87

RESULT 13
US-10-995-561-72930/c
; Sequence 72930, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 72930
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-72930

Query Match      64.0%; Score 16; DB 7; Length 201;
Best Local Similarity 79.2%; Pred. No. 1.5e+02;
Matches 19; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy  2 AAAAAGCAGACGAGCCCTCGG 25
Db  85 AAAAAGCAGACGAGCGGCACGCTG 62

RESULT 14
US-10-750-185-28365/c
; Sequence 28365, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
```

```
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 28365
; LENGTH: 1409
; TYPE: DNA
; ORGANISM: Bovine 19866881382047
US-10-750-185-28365

Query Match      64.0%; Score 16; DB 7; Length 1409;
Best Local Similarity 79.2%; Pred. No. 2.1e+02;
Matches 19; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy  2 AAAAAGCAGACGAGCCCTCGG 25
Db  64 AGAAAGAAACAGCAAGCCCTCGG 41

RESULT 15
US-10-750-623-28365/c
; Sequence 28365, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 28365
; LENGTH: 1409
; TYPE: DNA
; ORGANISM: Bovine 19866881382047
US-10-750-623-28365

Query Match      64.0%; Score 16; DB 7; Length 1409;
Best Local Similarity 79.2%; Pred. No. 2.1e+02;
Matches 19; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy  2 AAAAAGCAGACGAGCCCTCGG 25
Db  64 AGAAAGAAACAGCAAGCCCTCGG 41

RESULT 16
US-11-000-688-716
; Sequence 716, Application US/11000688
; Publication No. US20050287544A1
; GENERAL INFORMATION:
; APPLICANT: BERTUCCI, Francois
; APPLICANT: HOULGATTE, Remi
; APPLICANT: BIRNBAUM, Daniel
; TITLE OF INVENTION: GENE EXPRESSION PROFILING OF COLON CANCER WITH DNA ARRAYS
; FILE REFERENCE: 1423-R-03
; CURRENT APPLICATION NUMBER: US/11/000,688
; CURRENT FILING DATE: 2004-12-01
; PRIOR APPLICATION NUMBER: US 60/525,987
; PRIOR FILING DATE: 2003-12-01
; NUMBER OF SEQ ID NOS: 1596
; SOFTWARE: PatentIn version 3.2
```


; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Probe
US-11-136-527-352769

Query Match 63.2%; Score 15.8; DB 8; Length 25;
Best Local Similarity 89.5%; Pred. No. 1.4e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAACGACACAGCGAGC 19
||||| ||||| ||||| ||||| |||||
Db 25 AAAAAAGGACACAGCTAGC 7

RESULT 26

US-11-136-527-7464/c
; Sequence 7464, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 7464
; LENGTH: 1400
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-7464

Query Match 63.2%; Score 15.8; DB 8; Length 1400;
Best Local Similarity 89.5%; Pred. No. 2.6e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAACGACACAGCGAGC 19
||||| ||||| ||||| ||||| |||||
Db 480 AAAAAAGGACACAGCTAGC 462

RESULT 27

US-10-750-185-54749/c
; Sequence 54749, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: WM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 54749
; LENGTH: 1713
; TYPE: DNA
; ORGANISM: Bovine 19866880686743
US-10-750-185-54749

Query Match 63.2%; Score 15.8; DB 7; Length 1713;
Best Local Similarity 89.5%; Pred. No. 2.7e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAACGACACAGCGAGC 19
||||| ||||| ||||| ||||| |||||
Db 386 AAAAAAGGACACAGCAAGC 368

RESULT 28

US-10-750-623-54749/c
; Sequence 54749, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: WM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 54749
; LENGTH: 1713
; TYPE: DNA
; ORGANISM: Bovine 19866880686743
US-10-750-623-54749

Query Match 63.2%; Score 15.8; DB 7; Length 1713;
Best Local Similarity 89.5%; Pred. No. 2.7e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAACGACACAGCGAGC 19
||||| ||||| ||||| ||||| |||||
Db 386 AAAAAAGGACACAGCAAGC 368

RESULT 29

US-11-136-527-3368/c
; Sequence 3368, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3368
; LENGTH: 2383
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-3368

Query Match 63.2%; Score 15.8; DB 8; Length 2383;
Best Local Similarity 89.5%; Pred. No. 2.8e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAACGACACAGCGAGC 19
||||| ||||| ||||| ||||| |||||
Db 1463 AAAAAAGGACACAGCTAGC 1445

RESULT 30

US-10-995-561-62101


```
; SEQ ID NO 29381
; LENGTH: 794
; TYPE: DNA
; ORGANISM: Bovine 19866880796529
US-10-750-185-29381

Query Match          62.4%; Score 15.6; DB 7; Length 794;
Best Local Similarity 81.8%; Pred. No. 3e+02; 4; Indels 0; Gaps 0;
Matches 18; Conservative 0; Mismatches 0;

QY 1 AAAAAACGACACAGCGAGCCCT 22
   ||||| ||||| ||||| |||||
Db 618 AAAAAACACAAATCAAGCCCT 597

RESULT 35
US-10-750-623-29381/c
; Sequence 29381, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 29381
; LENGTH: 794
; TYPE: DNA
; ORGANISM: Bovine 19866880796529
US-10-750-623-29381

Query Match          62.4%; Score 15.6; DB 7; Length 794;
Best Local Similarity 81.8%; Pred. No. 3e+02; 4; Indels 0; Gaps 0;
Matches 18; Conservative 0; Mismatches 0;

QY 1 AAAAAACGACACAGCGAGCCCT 22
   ||||| ||||| ||||| |||||
Db 618 AAAAAACACAAATCAAGCCCT 597

RESULT 36
US-11-041-776-65/c
; Sequence 65, Application US/11041776
; Publication No. US20050272057A1
; GENERAL INFORMATION:
; APPLICANT: ABRAHAMSEN, MITCHELL
; APPLICANT: FREIJE, WADIAH
; TITLE OF INVENTION: SMALL SEGMENTS OF DNA DETERMINE ANIMAL IDENTITY AND
; FILE REFERENCE: 34579-97951
; CURRENT APPLICATION NUMBER: US/11/041,776
; CURRENT FILING DATE: 2005-01-24
; PRIOR APPLICATION NUMBER: 60/538,791
; PRIOR FILING DATE: 2004-01-23
; PRIOR APPLICATION NUMBER: 60/539,728
; PRIOR FILING DATE: 2004-01-26
; NUMBER OF SEQ ID NOS: 85
; SOFTWARE: Patentin Ver. 3.3
; SEQ ID NO 65
; LENGTH: 870
; TYPE: DNA
; ORGANISM: Sus scrofa
; FEATURE:
```

```
; NAME/KEY: misc feature
; LOCATION: (1) _ (870)
; OTHER INFORMATION: nucleotides 121-990 of LEPR
US-11-041-776-65

Query Match          62.4%; Score 15.6; DB 8; Length 870;
Best Local Similarity 81.8%; Pred. No. 3e+02; 4; Indels 0; Gaps 0;
Matches 18; Conservative 0; Mismatches 4;

QY 2 AAAAAACGACACAGCGAGCCCTC 23
   ||||| ||||| ||||| |||||
Db 528 AAAAAATGCACACGAAGTCATC 507

RESULT 37
US-10-750-185-59482
; Sequence 59482, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 59482
; LENGTH: 1045
; TYPE: DNA
; ORGANISM: Bovine 19866881274606
US-10-750-185-59482

Query Match          62.4%; Score 15.6; DB 7; Length 1045;
Best Local Similarity 81.8%; Pred. No. 3.1e+02; 4; Indels 0; Gaps 0;
Matches 18; Conservative 0; Mismatches 4;

QY 4 AAACGACACAGCGAGCCCTCGG 25
   ||||| ||||| ||||| |||||
Db 153 AGACGACAGCGGAGTCTCGG 174

RESULT 38
US-10-750-623-59482
; Sequence 59482, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 59482
; LENGTH: 1045
; TYPE: DNA
; ORGANISM: Bovine 19866881274606
```

US-10-750-623-59482

Query Match 62.4%; Score 15.6; DB 7; Length 1045;
Best Local Similarity 81.8%; Pred. No. 3.1e+02;
Matches 18; Conservative 0; Mismatches 4; Indels 0

Qy 4 AAACGACACAGCGAGCCCTCGG 25
153 AGACGACAGGGCGAGTCCTCGG 174
Db

RESULT 39

```

US-11-136-527-132
; Sequence 132, Application US/11136527
; Publication NO. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 132
; LENGTH: 1197
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-132

```

Query Match	62.4%	Score 15.6;	DB 8;	Length 1197;
Best Local Similarity	81.8%	Pred. No. 3.2e+02;		
Matches 18: Conservative	0;	Mismatches 4;	Indels 0	

Qy 2 AAAAAACGACACAGCGAGCCCTC 23
||||| ||| ||| ||| ||| |||
pb 307 AAAAAAGGAGACACCGGGCCCTC 328

RESULT 40

```

US-10-750-185-44586
; Sequence 44586, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: Denise, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 44586
; LENGTH: 1490
; TYPE: DNA
; ORGANISM: Bovine 19866880458485
US-10-750-185-44586

```

Query Match	62.4%;	Score 15.6;	DB 7;	Length 1490;
Best Local Similarity	81.8%;	Pred. No. 3.3e+02;		
Matches 18: Conservative	0;	Mismatches 4;	Indels 0	

QY 1 AAAAAAGACACAGCGAGCCCT 22

D_b 1397 AAAAATTGACACAGCAAGCTCT 1418

RESULT 41

```

US-10-750-623-44586
; Sequence 44586, Application US/10750623
; Publication No. US20050287531A1
;
; GENERAL INFORMATION:
;
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
;
; TITLE OF INVENTION: METHODS
; AND SYSTEMS FOR INFERRING BOVINE TRAITS
;
; FILE REFERENCE: MM1100-1
;
; CURRENT APPLICATION NUMBER: US/10750,623
; CURRENT FILING DATE: 2003-12-31
;
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
;
; NUMBER OF SEQ ID NOS: 64922
;
; SOFTWARE: PatentIN version 3.1
;
; SEQ. ID. NO 44586
;
; LENGTH: 1490
;
; TYPE: DNA
;
; ORGANISM: Bovine
;
US-10-750-623-44586

```

Query Match	62.4%	Score 15.6;	DB 7;	Length 1490;
Best Local Similarity	81.8%	Pred. No. 3.3e+02;		
Matches 18; Conservative	0;	Mismatches 4;	Indels 0;	Gaps 0;

QY 1 AAAAAACGACACAGCGAGCCT 22
 ||||| ||||| |||
db 1397 AAAAATTGACACAGCAAGCTCT 1418

RESULT 42

```

US-10-750-185-52468/C
; Sequence 52468, Application US/10750185
; Publication No. US200502603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 52468
; LENGTH: 1700
; TYPE: DNA
; ORGANISM: Bovine
US-10-750-185-52468

```

Query Match	62.4%	Score 15.6	DB 7	Length 1700
Best Local Similarity	81.8%	Pred. No. 3.4e+02		
Matches 18: Conservative	0	Mismatches 4	Indels 0	Gaps 0

QY 1 AAAAAAGACACAGCGGCCCT 22
||||| | ||||| |||||
DH 765 AAAAAAGAAAAACAGCAGGCCCT 744

```
RESULT 43
US-10-750-623-52468/c
; Sequence 52468, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 52468
; LENGTH: 1700
; TYPE: DNA
; ORGANISM: Bovine 19866881348662
US-10-750-623-52468

Query Match      62.4%; Score 15.6; DB 7; Length 1700;
Best Local Similarity 81.8%; Pred. No. 3.4e+02;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAACGACACGCGGCCCT 22
   ||||| ||||| ||||| |||||
DB 765 AAAAAAAGACAGAGGCCCT 744

RESULT 44
US-10-750-185-35903/c
; Sequence 35903, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 35903
; LENGTH: 1771
; TYPE: DNA
; ORGANISM: Bovine 19866881348662
US-10-750-185-35903

Query Match      62.4%; Score 15.6; DB 7; Length 1771;
Best Local Similarity 81.8%; Pred. No. 3.4e+02;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAACGACACGCGGCCCT 22
   ||||| ||||| ||||| |||||
DB 361 AAAAAAAGACAGAGGCCCT 340

RESULT 45
US-10-750-623-35903/c
; Sequence 35903, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 52468
; LENGTH: 1700
; TYPE: DNA
; ORGANISM: Bovine 19866881348662
US-10-750-623-35903

Query Match      62.4%; Score 15.6; DB 7; Length 1771;
Best Local Similarity 81.8%; Pred. No. 3.4e+02;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAACGACACGCGGCCCT 22
   ||||| ||||| ||||| |||||
DB 361 AAAAAAAGACAGAGGCCCT 340

RESULT 46
US-10-995-561-13357
; Sequence 13357, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13357
; LENGTH: 54767
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-13357

Query Match      62.4%; Score 15.6; DB 7; Length 54767;
Best Local Similarity 81.8%; Pred. No. 5.7e+02;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAACGACACGCGGCCCT 22
   ||||| ||||| ||||| |||||
DB 31689 AAAAAACACACAGCAAGCTCT 31710

RESULT 47
US-11-121-086-38
; Sequence 38, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; APPLICANT: NIELSEN, KIRSTEN V.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138.6000-00000
; CURRENT APPLICATION NUMBER: US/11/121,086
; CURRENT FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
```



```

; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 38
; LENGTH: 151169
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-38

```

```
Query Match      62.4%; Score 15.6; DB 8; Length 151169;
Best Local Similarity 81.8%; Pred. No. 6.5e+02;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
```

Qy 1 AAAAAACGACACAGCGAGCCCT 22
17421 AAAAAACGGCACGGCCAGCCAT 17442

RESULT 48

```

RESULT 48
US-10-517-151-105
; Sequence 105, Application US/10517151
; Publication No. US20060019252A1
; GENERAL INFORMATION:
; APPLICANT: Nakamura, Yusuke
; APPLICANT: Furukawa, Yoichi
; APPLICANT: Oncotherapy Science, Inc.
; APPLICANT: The University of Tokyo
; TITLE OF INVENTION: Genes and Polypeptides Relating to Hepatocellular or
; TITLE OF INVENTION: Colorectal Carcinoma
; FILE REFERENCE: 082379-000400US
; CURRENT APPLICATION NUMBER: US/10/517,151
; PRIOR FILING DATE: 2004-12-06
; PRIOR APPLICATION NUMBER: US 60/386,985
; PRIOR FILING DATE: 2002-06-06
; PRIOR APPLICATION NUMBER: WO PCT/JP03/07070
; PRIOR FILING DATE: 2003-06-04
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 105
; LENGTH: 55
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Artificially synthesized target sequence for siRNA
US-10-517-151-105

```

Query Match	61.6%;	Score 15.4;	DB 6;	Length 55;
Best Local Similarity	76.0%;	Pred. No. 2.4e+02;		
Matches 19:	Conservative	0;	Mismatches 6;	Indels 0;
				Gaps 0;

Qy 1 AAAAAACGACACAGCGAGCCCTCGG 25
|||||
pb 1 AAAAAACGAAACACCGATGACTGGG 25
|||||

RESULT 49

```

RESULT 49
US-10-310-914A-845/C
; Sequence 845, Application US/10310914A
; Publication No. US20060003322A1
; GENERAL INFORMATION:
; APPLICANT: Bentwich, Isaac
; APPLICANT: Shiler, Kvuza
; TITLE OF INVENTION: Bioinformatically detectable group of novel regulatory genes and
; FILE OF INVENTION: uses thereof
; FILE REFERENCE: 06087.0200.CPUS01
; CURRENT APPLICATION NUMBER: US/10/310,914A
; CURRENT FILING DATE: 2002-12-06
; NUMBER OF SEQ ID NOS: 1388402
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 845
; LENGTH: 121
; TYPE: RNA
; ORGANISM: Human

```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 3, 2006, 10:53:23 ; Search time 863 Seconds
(without alignments)
1646.682 Million cell updates/sec

Title: US-10-719-900-20

Perfect score: 25

Sequence: 1 aaaaactctgtaggcttcagtg 25

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database :

GenEmbl:

1: gb_ba.*

2: gb_in.*

3: gb_env.*

4: gb_ov.*

5: gb_pat.*

6: gb_ph.*

7: gb_pr.*

8: gb_ro.*

9: gb_sts.*

10: gb_sy.*

11: gb_un.*

12: gb_vi.*

13: gb_vt.*

14: gb_htg.*

15: gb_pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	25	100.0	74868	9 AF311213	AP311213 Mus muscu
C 2	25	100.0	178391	9 AC154141	AC154141 Mus muscu
C 3	20.8	83.2	128447	9 AC123864	AC123864 Mus muscu
C 4	20.8	83.2	194350	9 AC140328	AC140328 Mus muscu
C 5	20.2	80.8	97011	8 HS997015	AL035073 Human DNA
C 6	20.2	80.8	143254	5 BX649263	BX649263 Zebrafish
C 7	20.2	80.8	157481	8 AC025125	AC025125 Homo sapi
C 8	20.2	80.8	172000	14 AC024194	AC024194 Homo sapi
C 9	20.2	80.8	198136	14 AC090165	AC090165 Homo sapi
C 10	20.2	80.8	279675	14 AC161462	AC161462 Gallus ga
C 11	19.2	76.8	258	6 CQ662163	CQ662163 Sequence
C 12	19.2	76.8	289	6 CQ693741	CQ693741 Sequence
C 13	19.2	76.8	458	10 BV210905	BV210905 GSFT1_66
C 14	19.2	76.8	545	6 CQ923410	CQ923410 Sequence
C 15	19.2	76.8	562	6 CQ225863	CQ225863 Sequence
C 16	19.2	76.8	562	6 CQ300992	CQ300992 Sequence
C 17	19.2	76.8	562	6 CQ338275	CQ338275 Sequence
C 18	19.2	76.8	711	10 BV565426	BV565426 rcr67d10.

19	19.2	76.8	785	10	BV592074	BV592074
C 20	19.2	76.8	2079	9	AB003502	AB003502 Mus muscu
C 21	19.2	76.8	2167	9	BC028325	BC028325 Mus muscu
C 22	19.2	76.8	2359	9	BC031640	BC031640 Mus muscu
C 23	19.2	76.8	2523	8	BC009503	BC009503 Homo sapi
C 24	19.2	76.8	2527	6	AX833607	AX833607 Sequence
C 25	19.2	76.8	2527	8	AK095483	AK095483 Homo sapi
C 26	19.2	76.8	2586	6	CQ714266	CQ714266 Sequence
C 27	19.2	76.8	2587	6	CQ981521	CQ981521 Sequence
C 28	19.2	76.8	2587	6	CS030980	CS030980 Sequence
C 29	19.2	76.8	2587	6	CS036801	CS036801 Sequence
C 30	19.2	76.8	2587	6	CS039932	CS039932 Sequence
C 31	19.2	76.8	2587	6	CS045753	CS045753 Sequence
C 32	19.2	76.8	2587	6	CS086305	CS086305 Sequence
C 33	19.2	76.8	2587	6	HS737M10	HS737M10
C 34	19.2	76.8	2587	6	HS737M10	HS737M10
C 35	19.2	76.8	5570	6	CQ413058	CQ413058 Sequence
C 36	19.2	76.8	5570	6	CQ492962	CQ492962 Sequence
C 37	19.2	76.8	79158	5	AC160831	AC160831 Xenopus t
C 38	19.2	76.8	107172	8	HS737M10	HS737M10
C 39	19.2	76.8	115719	14	AC151438	AC151438 Carolinia
C 40	19.2	76.8	116273	14	AP000596	AP000596 Homo sapi
C 41	19.2	76.8	139376	8	HSU95742	HSU95742
C 42	19.2	76.8	148412	8	AC010654	AC010654 Homo sapi
C 43	19.2	76.8	150525	6	CS086332	CS086332 Sequence
C 44	19.2	76.8	157746	9	AC107677	AC107677 Mus muscu
C 45	19.2	76.8	159613	8	AP003420	AP003420 Homo sapi
C 46	19.2	76.8	160945	8	AP003420	AP003420 Homo sapi
C 47	19.2	76.8	161009	14	AC163997	AC163997 Mus muscu
C 48	19.2	76.8	163795	8	AP000356	AP000356 Homo sapi
C 49	19.2	76.8	163844	8	AC009048	AC009048 Homo sapi
C 50	19.2	76.8	163865	14	AC019081	AC019081 Homo sapi
C 51	19.2	76.8	165948	5	AC147169	AC147169 Xenopus t
C 52	19.2	76.8	166358	14	AC024184	AC024184 Homo sapi
C 53	19.2	76.8	166536	6	CS086331	CS086331 Sequence
C 54	19.2	76.8	168119	14	AC068686	AC068686 Homo sapi
C 55	19.2	76.8	168237	8	AC012622	AC012622 Homo sapi
C 56	19.2	76.8	169777	14	AC090778	AC090778 Homo sapi
C 57	19.2	76.8	170387	14	AC134388	AC134388 Pan trogl
C 58	19.2	76.8	170682	14	AC073256	AC073256 Homo sapi
C 59	19.2	76.8	170823	14	AC133549	AC133549 Homo sapi
C 60	19.2	76.8	171480	14	AC097333	AC097333 Pan trogl
C 61	19.2	76.8	172827	8	AC007216	AC007216 Homo sapi
C 62	19.2	76.8	173774	8	AL356495	AL356495 Human DNA
C 63	19.2	76.8	176731	5	AC147209	AC147209 Xenopus t
C 64	19.2	76.8	178727	14	AC149878	AC149878 Xenopus t
C 65	19.2	76.8	183271	8	AC104462	AC104462 Homo sapi
C 66	19.2	76.8	185404	14	AC025409	AC025409 Homo sapi
C 67	19.2	76.8	186299	14	AC155555	AC155555 Zea mays
C 68	19.2	76.8	186505	9	AC116394	AC116394 Mus muscu
C 69	19.2	76.8	192788	14	AC145062	AC145062 Pan trogl
C 70	19.2	76.8	194144	5	AC148460	AC148460 Xenopus t
C 71	19.2	76.8	200105	14	AC105308	AC105308 Pan trogl
C 72	19.2	76.8	204316	9	AC087541	AC087541 Mus muscu
C 73	19.2	76.8	216238	9	AL513356	AL513356 Mouse DNA
C 74	19.2	76.8	233387	14	AC162459	AC162459 Mus muscu
C 75	19.2	76.8	237308	14	AC009593	AC009593 Homo sapi
C 76	19.2	76.8	247254	14	AC114844	AC114844 Rattus no
C 77	18.8	75.2	23776	6	AR659661	AR659661 Sequence
C 78	18.8	75.2	57499	14	AC123371_4	Continuation (5 of
C 79	18.8	75.2	91186	14	AC118415_3	Continuation (4 of
C 80	18.8	75.2	97862	14	AC120823_4	Continuation (5 of
C 81	18.8	75.2	105741	14	AC140160	Rattus no
C 82	18.8	75.2	110000	14	AC140160_3	Continuation (4 of
C 83	18.8	75.2	110000	14	AC112426_2	Continuation (3 of
C 84	18.8	75.2	110000	14	AC115558_1	Continuation (2 of
C 85	18.8	75.2	114431	8	BX842590	BX842590 Rhesus DN
C 86	18.8	75.2	114493	14	AC142483	AC142483 Rattus no
C 87	18.8	75.2	114493	14	AC142483	AC142483 Rattus no
C 88	18.8	75.2	123897	14	AC144684	AC144684 Rattus no
C 89	18.8	75.2	134008	8	AC009892	AC009892 Homo sapi
C 90	18.8	75.2	142045	14	AC141650	AC141650 Rattus no
C 91	18.8	75.2	142579	14	AC136110	AC136110 Rattus no

c 92	18.8	75.2	148497	8	AL139117	AL139117 Human DNA
c 93	18.8	75.2	152134	5	BX571944	BX571944 Zebrafish
c 94	18.8	75.2	157487	14	AC142359	AC142359 Rattus no
c 95	18.8	75.2	165467	14	AC159033	AC159033 Pan trogl
c 96	18.8	75.2	177306	8	CR759950	CR759950 Gorilla D
c 97	18.8	75.2	183939	14	AC121650	AC121650 Rattus no
c 98	18.8	75.2	194020	9	AC121808	AC121808 Mus muscu
c 99	18.8	75.2	198817	9	AC124562	AC124562 Mus muscu
c 100	18.8	75.2	216124	14	AC098305	AC098305 Rattus no
c 101	18.8	75.2	219212	14	AC095817	AC095817 Rattus no
c 102	18.8	75.2	222939	14	AC111598	AC111598 Rattus no
c 103	18.8	75.2	229001	14	AC120233	AC120233 Rattus no
c 104	18.8	75.2	231670	14	AC142363	AC142363 Rattus no
c 105	18.8	75.2	233829	9	AC140397	AC140397 Mus muscu
c 106	18.8	75.2	237420	14	AC118158	AC118158 Rattus no
c 107	18.8	75.2	246043	14	AC105890	AC105890 Rattus no
c 108	18.8	75.2	247758	14	AC117041	AC117041 Rattus no
c 109	18.8	75.2	251166	14	AC163060	AC163060 Bos tauru
c 110	18.8	75.2	251589	14	AC111900	AC111900 Rattus no
c 111	18.8	75.2	254584	14	AC110948	AC110948 Rattus no
c 112	18.8	75.2	254732	14	AC111955	AC111955 Rattus no
c 113	18.8	75.2	262023	14	AC160648	AC160648 Pan trogl
c 114	18.8	75.2	262023	14	AC120890	AC120890 Rattus no
c 115	18.8	75.2	265731	14	AC120890	AC120890 Rattus no
c 116	18.8	75.2	265731	14	AC120890	AC120890 Rattus no
c 117	18.8	75.2	265833	14	AC125919	AC125919 Rattus no
c 118	18.8	75.2	270546	14	AC127017	AC127017 Rattus no
c 119	18.8	75.2	271418	14	AC113678	AC113678 Rattus no
c 120	18.8	75.2	276158	14	AC129229	AC129229 Rattus no
c 121	18.8	75.2	279971	14	AC103468	AC103468 Rattus no
c 122	18.8	75.2	284196	14	AC110447	AC110447 Rattus no
c 123	18.8	75.2	287138	14	AC114438	AC114438 Rattus no
c 124	18.8	75.2	291031	14	AC117885	AC117885 Rattus no
c 125	18.8	75.2	292888	14	AC095400	AC095400 Rattus no
c 126	18.8	75.2	295460	14	AC114116	AC114116 Rattus no
c 127	18.8	75.2	316118	14	AC112820	AC112820 Rattus no
c 128	18.8	75.2	349750	14	AC109776	AC109776 Rattus no
c 129	18.6	74.4	609	10	BV001701	BV001701 S208P6127
c 130	18.6	74.4	6043	5	BC070842	BC070842 Xenopus l
c 131	18.6	74.4	27998	8	AC128713	AC128713 Homo sapi
c 132	18.6	74.4	38968	8	AC000075	AC000075 Homo sapi
c 133	18.6	74.4	39982	6	AR308846	AR308846 Sequence
c 134	18.6	74.4	39982	6	AR580265	AR580265 Sequence
c 135	18.6	74.4	54569	9	BX571685	BX571685 Mouse DNA
c 136	18.6	74.4	63753	14	AC090089	AC090089 Homo sapi
c 137	18.6	74.4	87849	2	AC084453	AC084453 Caenorhab
c 138	18.6	74.4	99486	8	AC090082	AC090082 Homo sapi
c 139	18.6	74.4	110000	15	AE016819_06	Continuation (7 of
c 140	18.6	74.4	119346	14	AC164752	Continuation (8 of
c 141	18.6	74.4	130278	8	AC005664	AC164752 Bos tauru
c 142	18.6	74.4	132634	14	AC123394	AC005664 Homo sapi
c 143	18.6	74.4	135245	8	AC073907	AC123394 Rattus no
c 144	18.6	74.4	142439	14	AC147432	AC073907 Homo sapi
c 145	18.6	74.4	162047	5	BX784036	AC147432 Pan trogl
c 146	18.6	74.4	162269	8	AC000097	BX784036 Zebrafish
c 147	18.6	74.4	163028	9	AC127258	AC000097 Homo sapi
c 148	18.6	74.4	163050	14	AC136176	AC127258 Mus muscu
c 149	18.6	74.4	163865	14	AC019081	AC136176 Rattus no
c 150	18.6	74.4	165269	8	AC140951	AC019081 Homo sapi
						AC140951 Pan trogl

ALIGNMENTS

RESULT 1	AF311213	74868 bp	DNA	linear	ROD 20-NOV-2002
LOCUS	Mus musculus BAC 171m12 MESDC1 (Mesdc1) and MESDC2 (Mesdc2) genes,				
DEFINITION	complete cds.				
ACCESSION	AF311213				
VERSION	AF311213.1	GI:11245449			
KEYWORDS	Mus musculus (house mouse)				
SOURCE					

ORGANISM	Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Mus.
AUTHORS	1 (bases 1 to 74868) Wines,M.E., Lee,L., Katari,M.S., Zhang,L., DeRossi,C., Shi,Y., Perkins,S., Feldman,M., McCombie,W.R. and Holdener,B.C.
TITLE	Identification of mesoderm development (mesd) candidate genes by comparative mapping and genome sequence analysis
JOURNAL	Genomics 72 (1), 88-98 (2001)
REFERENCE	2 (bases 1 to 74868) Perkins,S., Feldman,M., McCombie,W.R. and Holdener,B.C.
AUTHORS	Direct Submission Submitted (27-SEP-2000) Biochemistry and Cell Biology, SUNY at Stony Brook, 346 Center for Molecular Medicine, Stony Brook, NY 11794-5215, USA
TITLE	Wines,M.E. Dr. Lee,L., Katari,M.S., Zhang,L., DeRossi,C., Shi,Y., Perkins,S., Feldman,M., McCombie,W.R. and Holdener,B.C.
JOURNAL	Direct Submission Submitted (27-SEP-2000) Biochemistry and Cell Biology, SUNY at Stony Brook, 346 Center for Molecular Medicine, Stony Brook, NY 11794-5215, USA
FEATURES	Location/Qualifiers
source	1..74868 /organism="Mus musculus" /mol_type="genomic DNA" /strain="129/SvJ" /db_xref="taxon:10090" /clone="BAC 171m12"
repeat_region	1207..1421 /note="RepeatMasker: B3" /rpt_family="SINE/B2" /rpt_type="dispersed"
repeat_region	complement(1516..1669) /note="RepeatMasker: B1F" /rpt_family="SINE/Alu"
repeat_region	/rpt_type="dispersed" complement(1815..1961) /note="RepeatMasker: B1_MM" /rpt_family="SINE/Alu"
repeat_region	/rpt_type="dispersed" 3013..3100 /note="RepeatMasker: MSTC" /rpt_family="LTR/MaLR"
repeat_region	/rpt_type="dispersed" 3230..3333 /note="RepeatMasker: MT2A" /rpt_family="LTR/MaLR"
repeat_region	/rpt_type="dispersed" 3334..3824 /note="RepeatMasker: RMER19B" /rpt_family="LTR"
repeat_region	/rpt_type="dispersed" complement(3876..4066) /note="RepeatMasker: ORR1C"
repeat_region	/rpt_family="LTR/MaLR" /rpt_type="dispersed" complement(4923..5029) /note="RepeatMasker: L3"
repeat_region	/rpt_family="LINE/CR1" /rpt_type="dispersed" complement(5083..5152) /note="RepeatMasker: L2"
repeat_region	/rpt_family="LINE/L2" /rpt_type="dispersed" complement(6386..6675) /note="RepeatMasker: B4"
repeat_region	/rpt_family="SINE/B4" /rpt_type="dispersed" complement(6817..7147) /note="RepeatMasker: MTE"
repeat_region	/rpt_family="LTR/MaLR" /rpt_type="dispersed" complement(7364..7474) /note="RepeatMasker: ID5"
repeat_region	/rpt_family="SINE/ID" /rpt_type="dispersed"

```

repeat_region 7869..7979
/note="RepeatMasker: B4"
/rpt_family="SINE/B4"
/rpt_type=dispersed
repeat_region complement(8175..8252)
/note="RepeatMasker: MIR"
/rpt_family="SINE/MIR"
/rpt_type=dispersed
repeat_region complement(8388..8590)
/note="RepeatMasker: B3"
/rpt_family="SINE/B2"
/rpt_type=dispersed
repeat_region complement(8595..8713)
/note="RepeatMasker: B1F"
/rpt_family="SINE/Alu"
/rpt_type=dispersed
repeat_region complement(8728..8952)
/note="RepeatMasker: B3A"
/rpt_family="SINE/B2"
/rpt_type=dispersed
repeat_region 10428..10504
/note="RepeatMasker: RMER15"
/rpt_family="LTR"
/rpt_type=dispersed
repeat_region 10681..10812
/note="RepeatMasker: RSINE1"
/rpt_family="SINE/B4"
/rpt_type=dispersed
repeat_region 11110..11254
/note="RepeatMasker: B1F"
/rpt_family="SINE/Alu"
/rpt_type=dispersed
repeat_region complement(12197..12336)
/note="RepeatMasker: RSINE1"
/rpt_family="SINE/B4"
/rpt_type=dispersed
repeat_region 12593..13088
/note="RepeatMasker: LIMA4"
/rpt_family="LINE/L1"
/rpt_type=dispersed
repeat_region 13168..13341
/note="RepeatMasker: B2_Mm1"
/rpt_family="SINE/B2"
/rpt_type=dispersed
repeat_region 13378..13432
/note="RepeatMasker: B2_Mm2"
/rpt_family="SINE/B2"
/rpt_type=dispersed
repeat_region complement(13952..14145)
/note="RepeatMasker: URR1A"
/rpt_family="DNA/MER1_type"
/rpt_type=dispersed
repeat_region 14384..14477
/note="RepeatMasker: MIR"
/rpt_family="SINE/MIR"
/rpt_type=dispersed
repeat_region 15620..15747
/note="RepeatMasker: RSINE1"
/rpt_family="SINE/B4"
/rpt_type=dispersed
repeat_region 16036..16115
/note="RepeatMasker: MIR"
/rpt_family="SINE/MIR"
/rpt_type=dispersed
repeat_region 16237..16368
/note="RepeatMasker: B1F"
/rpt_family="SINE/Alu"
/rpt_type=dispersed
repeat_region complement(16622..16763)
/note="RepeatMasker: MIR3"
/rpt_family="SINE/MIR"
/rpt_type=dispersed
repeat_region 17031..17087

```

```

/note="RepeatMasker: PB1D10"
/rpt_family="SINE/Alu"
/rpt_type=dispersed
complement(18601..18893)
/note="RepeatMasker: B4A"
/rpt_family="SINE/B4"
/rpt_type=dispersed
19072..19258
/note="RepeatMasker: B3"
/rpt_family="SINE/B2"
/rpt_type=dispersed
complement(19709..19886)
/note="RepeatMasker: B4A"
/rpt_family="SINE/B4"
/rpt_type=dispersed
20177..20657
/note="RepeatMasker: RMER4"
/rpt_family="LTR"
/rpt_type=dispersed
21857..22102
/note="RepeatMasker: RMER15"
/rpt_family="LTR"
/rpt_type=dispersed
22456..22585
/note="RepeatMasker: B1_MM"
/rpt_family="SINE/Alu"
/rpt_type=dispersed
complement(22971..23063)
/note="RepeatMasker: RMER15"
/rpt_family="LTR"
/rpt_type=dispersed
23529..23653
/note="RepeatMasker: RSINE1"
/rpt_family="SINE/B4"
/rpt_type=dispersed
23690..23707
/note="RepeatMasker: RSINE1"
/rpt_family="SINE/B4"
/rpt_type=dispersed
complement(23733..23870)
/note="RepeatMasker: B1_MM"
/rpt_family="SINE/Alu"
/rpt_type=dispersed
24067..24243
/note="RepeatMasker: B3"
/rpt_family="SINE/B2"
/rpt_type=dispersed

```

Query Match 100.0%; Score 25; DB 9; Length 74868;
Best Local Similarity 100.0%; Pred. No. 0.14;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 AAAAACTCTGTAGGCTTTCAGTG 25
    |||||
Db 33482 AAAAACTCTGTAGGCTTTCAGTG 33458

```

```

RESULT 2
AC154141 178391 bp DNA linear ROD 21-JUN-2005
LOCUS Mus musculus chromosome 7, clone RP24-343G19, complete sequence.
DEFINITION
AC154141
ACCESSION AC154141.7 GI:68051863
VERSION HTG.
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 178391)
AUTHORS Birren,B., Nusbaum,C. and Lander,E.
TITLE Mus musculus chromosome 7, clone RP24-343G19
JOURNAL Unpublished

```

REFERENCE
AUTHORS

2 (bases 1 to 178391)
 Birren,B., Nussbaum,C., Lander,E., Abouelleil,A., Allen,N.,
 Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V.,
 Bloom,T., Boguslavskiy,L., Boukhgalter,B., Camarata,J., Chang,J.,
 Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B.,
 DeArellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L.,
 Erickson,J., Faro,S., Ferreira,P., Fitzgerald,M., Gage,D.,
 Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N.,
 Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I.,
 Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T.,
 Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbitt,R.,
 MacLean,C., Macdonald,P., Major,J., Manning,J., Matthews,C.,
 McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
 Murphy,T., Naylor,J., Nguyen,C., Nguyen,T., Nicol,R., Norbu,C.,
 O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,
 Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C.,
 Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupback,R.,
 Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
 Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J.,
 Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R.,
 Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L.,
 Zimmer,A. and Zody,M.

TITLE
JOURNAL

Submitted (24-DEC-2004) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE
AUTHORS

3 (bases 1 to 178391)
 Birren,B., Nussbaum,C., Lander,E., Abouelleil,A., Allen,N.,
 Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V.,
 Bloom,T., Boguslavskiy,L., Boukhgalter,B., Camarata,J., Chang,J.,
 Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B.,
 DeArellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L.,
 Erickson,J., Faro,S., Ferreira,P., Fitzgerald,M., Gage,D.,
 Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N.,
 Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I.,
 Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T.,
 Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbitt,R.,
 MacLean,C., Macdonald,P., Major,J., Manning,J., Matthews,C.,
 McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
 Murphy,T., Naylor,J., Nguyen,C., Nguyen,T., Nicol,R., Norbu,C.,
 O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,
 Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C.,
 Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupback,R.,
 Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
 Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J.,
 Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R.,
 Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L.,
 Zimmer,A. and Zody,M.

TITLE
JOURNAL

Submitted (28-APR-2005) Broad Institute of MIT and Harvard, 320
 Charles Street, Cambridge, MA 02141, USA

REFERENCE
AUTHORS

4 (bases 1 to 178391)
 Birren,B., Nussbaum,C., Lander,E., Abouelleil,A., Allen,N.,
 Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V.,
 Bloom,T., Boguslavskiy,L., Boukhgalter,B., Camarata,J., Chang,J.,
 Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B.,
 DeArellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L.,
 Erickson,J., Faro,S., Ferreira,P., Fitzgerald,M., Gage,D.,
 Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N.,
 Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I.,
 Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T.,
 Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbitt,R.,
 MacLean,C., Macdonald,P., Major,J., Manning,J., Matthews,C.,
 McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
 Murphy,T., Naylor,J., Nguyen,C., Nguyen,T., Nicol,R., Norbu,C.,
 O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,
 Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C.,
 Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupback,R.,
 Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
 Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J.,
 Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R.,
 Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L.,
 Zimmer,A. and Zody,M.

TITLE

Submitted (21-JUN-2005) Broad Institute of MIT and Harvard, 320
 Charles Street, Cambridge, MA 02141, USA

JOURNAL

COMMENT

On Jun 21, 2005 this sequence version replaced gi:62945640.
 All repeats were identified using RepeatMasker.
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Broad Institute of MIT and Harvard

Center code: W1BR

Web site: http://www-seq.wi.mit.edu

Contact: sequence_submissions@broad.mit.edu

----- Project Information

Center project name: L30897

Center clone name: 343_G_19

FEATURES

source

Location/Qualifiers

1..178391
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /chromosome="7"
 /map="7"
 /clone="RP24-343G19"
 /clone_lib="RPCI-24 Male Mouse BAC"
 1585..1757
 /rpt_family="GA-rich"
 complement(2018..2295)
 /rpt_family="MLTIG1"
 2324..2371
 /rpt_family="GA-rich"
 2517..2624
 /rpt_family="PB1D7"
 complement(3040..3099)
 /rpt_family="MTC"
 4765..5049
 /rpt_family="B4"
 6950..7018
 /rpt_family="CTGGGGn"
 7195..7229
 /rpt_family="TG)n"
 7258..7451
 /rpt_family="ID_B1"
 9551..9639
 /rpt_family="CT-rich"
 9925..9992
 /rpt_family="MIR"
 complement(10082..10496)
 /rpt_family="MTC"
 10634..10684
 /rpt_family="TG)n"
 10684..10745
 /rpt_family="GA)n"
 10747..10819
 /rpt_family="TCTG)n"
 complement(10821..11112)
 /rpt_family="B4"
 12146..12231
 /rpt_family="GA-rich"
 12244..12428
 /rpt_family="GGGAGA)n"
 14033..14105
 /rpt_family="TC)n"
 14105..14170
 /rpt_family="TTCC)n"
 14433..14611
 /rpt_family="L2"
 complement(14959..15592)
 /rpt_family="L1MB8"
 complement(16298..16382)
 /rpt_family="B1F"
 16740..16858
 /rpt_family="L1_MM"
 16859..16915

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region


```
repeat_region /rpt_family="B4"
1590_ .1774
/rpt_family="B2"
1809_ .1863
/rpt_family="B4"
1864_ .2215
/rpt_family="MaLR"
2216_ .2362
/rpt_family="Alu"
2363_ .2394
/rpt_family="MaLR"
2408_ .2499
/rpt_family="B4"
2446_ .2517
/rpt_family="tRNA-Ala-GCY_"
2604_ .2750
/rpt_family="Alu"
3000_ .3080
/rpt_family="B4"
3081_ .3197
/rpt_family="Alu"
3270_ .3425
/rpt_family="B4"
3426_ .3540
/rpt_family="B4"
3580_ .3650
/rpt_family="B4"
3698_ .4076
/rpt_family="MaLR"
4639_ .4832
/rpt_family="B4"
4911_ .5013
/rpt_family="B4"
6445_ .6610
/rpt_family="B2"
6671_ .6769
/rpt_family="Alu"
7257_ .7449
/rpt_family="B2"
7545_ .7736
/rpt_family="B2"
7551_ .7624
/product="tRNA-Ser"
/notes="Likely pseudogene (HMM Sc=40.07 / Sec struct
Sc=-13.49)"
repeat_region 7784_ .7955
/rpt_family="B2"
8011_ .8158
/rpt_family="L1"
8161_ .8471
/rpt_family="MER1_type"
8588_ .8833
/rpt_family="B4"
9003_ .9082
/rpt_family="MER1_type"
9990_ .10129
/rpt_family="Alu"
9998_ .10170
/rpt_family="B4"
10451_ .10585
/rpt_family="Alu"
11281_ .11356
/rpt_family="ERV1"
12025_ .12214
/rpt_family="B2"
12217_ .12344
/rpt_family="Alu"
12498_ .12636
/rpt_family="Alu"
12681_ .12835
/rpt_family="B4"
13289_ .13498
/rpt_family="B2"
```

```
repeat_region 13600_ .13676
/rpt_family="Alu"
13731_ .13839
/rpt_family="Alu"
13984_ .14128
/rpt_family="Alu"
14135_ .14239
/rpt_family="B4"
14416_ .14590
/rpt_family="B4"
14622_ .14801
/rpt_family="B2"
14817_ .14963
/rpt_family="Alu"
15183_ .15373
/rpt_family="B2"
15416_ .15567
/rpt_family="B4"
15565_ .15663
/rpt_family="B4"
15714_ .15860
/rpt_family="Alu"
16241_ .16338
/rpt_family="BC1_MM"
16425_ .16554
/rpt_family="MER1_type"
16589_ .16789
/rpt_family="B2"
17586_ .17724
/rpt_family="B4"
17741_ .17922
/rpt_family="B2"
17937_ .17994
```

Query Match 83.2%; Score 20.8; DB 9; Length 128447;

Best Local Similarity 91.7%; Pred. No. 22;

Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAACTCTGTAGCTTCAGT 24

DB 90615 AAAAACTCTGTAGCTTCAGT 90638

RESULT 4
AC140328/c AC140328 194350 bp DNA linear ROD 27-JAN-2005
LOCUS Mus musculus BAC clone RP23-404L15 from 8, complete sequence.
DEFINITION AC140328
ACCESSION AC140328
VERSION AC140328.3 GI:51854628
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE
AUTHORS Waligorski,J., Kozlowski,A. and Haglund,K.
TITLE The sequence of Mus musculus BAC clone RP23-404L15
JOURNAL Unpublished (2001)
REFERENCE 2 (bases 1 to 194350)
AUTHORS McPherson,J.D. and Waterston,R.H.
TITLE Direct Submission

JOURNAL Submitted (23-FEB-2003) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
REFERENCE 3 (bases 1 to 194350)
AUTHORS Wilson,R.K.
TITLE Direct Submission

JOURNAL Submitted (22-JUL-2004) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
REFERENCE 4 (bases 1 to 194350)
AUTHORS Wilson,R.K.
TITLE Direct Submission

JOURNAL Submitted (02-SEP-2004) Genome Sequencing Center, 4444 Forest Park

REFERENCE AUTHORS TITLE JOURNAL

Parkway, St. Louis, MO 63108, USA
5 (bases 1 to 194350)
Wilson, R.K.
Direct Submission
Submitted (27-JAN-2005) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
On Sep 2, 2004 this sequence version replaced gi:50511686.

COMMENT

----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu>
Contact: submissions@wustl.edu
----- Summary Statistics
Center project name: M_BA0404L15

NOTICE:

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by
restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. Wes Warren,
Department of Genetics, Washington University, St. Louis MO. For
additional information about the map position of this sequence, see
<http://genome.wustl.edu>

SOURCE INFORMATION:

The RPCI-23 BAC Library has been constructed by Kazutoyo Osegawa
and Minako Tatenno in the laboratory of Pieter de Jong
(<http://www.chori.org>) from female C57BL/6J mouse kidney and/or
brain genomic DNA. The clone and detailed information can be
obtained from Research Genetics, Inc. (<http://www.resgen.com>) or
Pieter de Jong and coworkers at <http://www.chori.org>

NEIGHBORING SEQUENCE INFORMATION:

This sequence is the entire insert of the clone.

FEATURES

source

1. .194350
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="8"
/map="8"

/clone="RP23-404L15"
/clone_lib="RPCI-23"

257. .346

repeat_region

/rpt_family="B4"

repeat_region

383. .518

/rpt_family="B4"

repeat_region

1111. .1252

/rpt_family="MIR"

repeat_region

1303. .1563

/rpt_family="B4"

repeat_region

1574. .1710

/rpt_family="Alu"

repeat_region

1952. .2091

/rpt_family="B4"

repeat_region

2783. .2866

/rpt_family="B4"

repeat_region

3025. .3143

/rpt_family="Alu"

repeat_region

3217. .3320

/rpt_family="Alu"

repeat_region

3462. .3621

/rpt_family="B2"

repeat_region

3726. .3862

repeat_region
3915. .4120
/rpt_family="B2"
repeat_region
4160. .4295
/rpt_family="Alu"
repeat_region
4575. .5085
/rpt_family="L1"
repeat_region
5359. .5455
/rpt_family="B4"
repeat_region
5479. .5519
/rpt_family="B4"
repeat_region
5520. .5618
/rpt_family="Alu"
repeat_region
5807. .5960
/rpt_family="B4"
repeat_region
5983. .6133
/rpt_family="Alu"
repeat_region
6144. .6189
/rpt_family="L1"
repeat_region
6324. .6437
/rpt_family="Alu"
repeat_region
6570. .6709
/rpt_family="B2"
repeat_region
6775. .6919
/rpt_family="Alu"
repeat_region
6942. .7134
/rpt_family="L1"
repeat_region
8118. .8523
/rpt_family="MaLR"
repeat_region
8582. .8720
/rpt_family="Alu"
repeat_region
8923. .9152
/rpt_family="MaLR"
repeat_region
10037. .10170
/rpt_family="Alu"
repeat_region
10191. .10337
/rpt_family="L2"
repeat_region
11563. .11692
/rpt_family="B4"
repeat_region
12166. .12289
/rpt_family="Alu"
repeat_region
12748. .12873
/rpt_family="Alu"
repeat_region
12874. .12952
/rpt_family="ID"
12878. .12948
/product="tRNA-Phe"
/note="Likely pseudogene (HMM Sc=23.61 / Sec struct
Sc=-0.14)"
repeat_region
12953. .12963
/rpt_family="Alu"
repeat_region
13038. .13150
/rpt_family="Alu"
repeat_region
13882. .13984
/rpt_family="Alu"
repeat_region
14084. .14237
/rpt_family="L1"
repeat_region
14504. .14907
/rpt_family="L1"
repeat_region
14908. .15075
/rpt_family="L1"
repeat_region
15084. .15206
/rpt_family="Alu"
repeat_region
15223. .15377
/rpt_family="B2"
repeat_region
15377. .15471
/rpt_family="Alu"
repeat_region
15996. .16132
/rpt_family="B4"
repeat_region
16162. .16310
/rpt_family="Alu"
16470. .16520
/note="Sequence derived from one plasmid subclone."
unare

```
repeat_region 16530..16667
/rpt_family="Alu"
repeat_region 16706..16746
/rpt_family="Alu"
repeat_region 16828..17037
/rpt_family="B2"
misc_feature 17033..17063
/note="Sequence derived from PCR product of genomic DNA."
repeat_region 17048..17194
/rpt_family="Alu"
repeat_region 17296..17367
/rpt_family="Alu"
repeat_region 17669..17906
/rpt_family="ERVK"
repeat_region 17907..18035
/rpt_family="Alu"
repeat_region 18054..18351
/rpt_family="B4"
repeat_region 18595..18738
/rpt_family="Alu"
repeat_region 18813..18884
/rpt_family="BC1_MM"
repeat_region 18829..18886
/rpt_family="tRNA-Ala-GCY_"
repeat_region 19023..19150
/rpt_family="B4"
```

```
Query Match 83.2%; Score 20.8; DB 9; Length 194350;
Best Local Similarity 91.7%; Pred. No. 22;
Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 1 AAAAACTCTGTAGCTTTTCAGT 24
|||||
Db 148168 AAAAACTCTGTAGCTTTACAGT 148145
```

```
RESULT 5
HS997015/c
LOCUS HS997015 97011 bp DNA linear PRI 18-MAY-2005
DEFINITION Human DNA sequence from clone RPS-997015 on chromosome 20p11.21-12
Contains genomic marker D20S904, STSs and GSSs, complete sequence.
ACCESSION AL035073
VERSION AL035073.7 GI:11493139
KEYWORDS HTG; D20S904.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE 1 (bases 1 to 97011)
AUTHORS Phillimore,B.
TITLE Direct Submission
JOURNAL Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vegas@sanger.ac.uk
Clone requests: clonerequest@sanger.ac.uk
```

```
On Nov 30, 2000 this sequence version replaced gi:9795193.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em.; EMBL; Sw.; SWISSPROT; Tr.; TREMBL; Wp.; WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep
This sequence was generated from part of bacterial clone contigs of human
chromosome 20, constructed by the Sanger Centre Chromosome 20
Mapping Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr20
RPS-997015 is from the library RPCI-5 constructed by the group of
Pieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pCYPAC2
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
```

Contact: vegas@sanger.ac.uk

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

FEATURES

```
source
location/Qualifiers
1..97011
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="20"
/map="p11.21-12"
/clone="RPS-997015"
/clone_lib="RPCI-5"
```

```
misc_feature 1
/note="Clone_left_end: RPS-997015"
```

```
misc_feature 97011
/note="Clone_right_end: RPS-997015"
```

ORIGIN

```
Query Match 80.8%; Score 20.2; DB 8; Length 97011;
Best Local Similarity 88.0%; Pred. No. 45;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 1 AAAAACTCTGTAGCTTTTCAGT 25
```

```
Db 54155 AGAAAACTCTGTAGGCATTCACGT 54131
```

```
RESULT 6
BX649263/c
LOCUS BX649263 143254 bp DNA linear VRF 02-DEC-2004
DEFINITION Zebrafish DNA sequence from clone DKEY-21022 in linkage group 4
Contains two novel genes, complete sequence.
ACCESSION BX649263
VERSION BX649263.6 GI:42415036
KEYWORDS HTG.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
REFERENCE 1 (bases 1 to 143254)
AUTHORS Pelan,S.
TITLE Direct Submission
JOURNAL Submitted (02-DEC-2004) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
zfish-help@sanger.ac.uk Clone requests:
http://www.sanger.ac.uk/Projects/D_rerio/fags.shtml#dataeight
On Feb 5, 2004 this sequence version replaced gi:41392404.
```

COMMENT

----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: zfish-help@sanger.ac.uk

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:
Em.; EMBL; Sw.; SWISSPROT; Tr.; TREMBL; Wp.; WORMPEP; Information on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep
Clone-derived Zebrafish pUC subclones occasionally display inconsistency over the length of mononucleotide A/T runs and conserved TA repeats. Where this is found the longest good quality representation will be submitted.
DKEY-21022 is from a Zebrafish BAC library
VECTOR: pIndigoBAC-5
This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate

SEQUENCE, 13 unordered pieces.
AC024194
AC024194.2 GI:8099107
HTG: HTGS PHASE1; HTGS_DRAFT.
SOURCE
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominae; Homo.
REFERENCE
1 (bases 1 to 172000)
Waterston,R.H.
The sequence of Homo sapiens clone
Unpublished
REFERENCE
2 (bases 1 to 172000)
Waterston,R.H.
Direct Submission
Submitted (25-FEB-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
COMMENT
On May 26, 2000 this sequence version replaced gi:7105580.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu/gsc/index.shtml>
----- Project Information -----
Center project name: H NH0587F22
----- Summary Statistics -----
Sequencing vector: M13, 100%
Sequencing method: plasmid, 0%
Chemistry: Dye-terminator Big Dye, 0% of reads
Assembly: Dye-terminator Big Dye, 0% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 164472 bases at least Q40
Consensus quality: 166662 bases at least Q30
Consensus quality: 167979 bases at least Q20
Insert size: 186000; agarose-fp
Insert size: 170800; sum-of-contigs
Quality coverage: 4.61 in Q20 bases; agarose-fp
Quality coverage: 5.07 in Q20 bases; sum-of-contigs

----- NOTE: This is a 'working draft' sequence. It currently consists of 13 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 1310: contig of 1310 bp in length
1311: gap of unknown length
1411: contig of 1385 bp in length
2796: gap of unknown length
2896: contig of 1071 bp in length
3967: gap of unknown length
4067: contig of 1615 bp in length
5681: gap of unknown length
5781: contig of 1797 bp in length
7578: gap of unknown length
7679: contig of 3398 bp in length
11077: gap of unknown length
21069: contig of 9892 bp in length
21169: gap of unknown length
35302: contig of 14133 bp in length
35402: gap of unknown length
35402: contig of 13019 bp in length
48421: gap of unknown length
48521: contig of 12610 bp in length
61131: gap of unknown length
61231: gap of unknown length
93885: contig of 32654 bp in length
93985: gap of unknown length
125773: contig of 31788 bp in length
125873: gap of unknown length

* 125873 172000: contig of 46128 bp in length.
FEATURES
source
1. .172000
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="17"
/clone="RP11-587F22"
1. .1310
/note="assembly_name:Contig45"
1311. .1410
/estimated_length=unknown
1411. .2795
/note="assembly_name:Contig48"
2796. .2895
/estimated_length=unknown
2896. .3966
/note="assembly_name:Contig49"
3967. .4066
/estimated_length=unknown
4067. .5681
/note="assembly_name:Contig58"
5682. .5781
/estimated_length=unknown
5782. .7578
/note="assembly_name:Contig59"
clone_end:T7
vector_side:right
7579. .7678
/estimated_length=unknown
7679. .11076
/note="assembly_name:Contig60"
clone_end:SP6
vector_side:right
11077. .11176
/estimated_length=unknown
11177. .21068
/note="assembly_name:Contig61"
21069. .21168
/estimated_length=unknown
21169. .35301
/note="assembly_name:Contig62"
35302. .35401
/estimated_length=unknown
35402. .48420
/note="assembly_name:Contig63"
48421. .48520
/estimated_length=unknown
48521. .61130
/note="assembly_name:Contig64"
61131. .61230
/estimated_length=unknown
61231. .93884
/note="assembly_name:Contig65"
93885. .93984
/estimated_length=unknown
93985. .125772
/note="assembly_name:Contig66"
125773. .125872
/estimated_length=unknown
125873. .172000
/note="assembly_name:Contig67"

ORIGIN

Query Match 80.8%; Score 20.2; DB 14; Length 172000;
Best Local Similarity 88.0%; Pred. No. 45;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AAAAACTCTGTAGGCTTTTCAGTG 25
|||||

Db 45841 AAAAACTCTGTAGGCTTTTCAGTG 45817
|||||

```

RESULT 9
AC090165/c
LOCUS
DEFINITION
    Homo sapiens chromosome 17 clone RP11-587F22 map 17, WORKING DRAFT
    SEQUENCE, 32 unordered pieces.
ACCESSION
    AC090165
VERSION
    AC090165.2 GI:14210559
KEYWORDS
    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE
    Homo sapiens
ORGANISM
    Homo sapiens
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
    Homidae; Homo.
REFERENCE
    1 (bases 1 to 198136)
    Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,
    Barna,N., Bastien,V., Boguslavsky,I., Boukhalter,B., Brown,A.,
    Camarata,J., Campopiano,A., Choepel,Y., Colangelo,M., Collins,S.,
    Collymore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S.,
    Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J.,
    Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N.,
    Haggis,B., Hearford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
    Jones,C., Karatas,A., Lakoque,K., Lamazares,R., Landers,T.,
    Lehoczyk,J., Levine,R., Liu,G., MacLean,C., Macdonald,P.,
    Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K.,
    McPheeters,R., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
    Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H.,
    O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
    Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R.,
    Riback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M.,
    Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S., Severy,P.,
    Sounez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
    Strauss,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J.,
    Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A.,
    Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J.,
    Zembek,L., Zimmer,A. and Zody,N.
    Direct Submission
    Submitted (17-FEB-2001) Whitehead Institute/MIT Center for Genome
    Research, 320 Charles Street, Cambridge, MA 02141, USA
    On May 27, 2001 this sequence version replaced gi:12957795.
    All repeats were identified using RepeatMasker:
    Smit, A.F.A. & Green, P. (1996-1997)
    http://ftp.genome.washington.edu/RM/RepeatMasker.html
    ----- Genome Center
    Center: Whitehead Institute/ MIT Center for Genome Research
    Center code: WIER
    Web site: http://www-seq.wi.mit.edu
    Contact: sequence_submissions@genome.wi.mit.edu
    ----- Project Information
    Center project name: L12036
    Center clone name: 587.F.22
    ----- Summary Statistics
    Sequencing vector: Plasmid; n/a; 100% of reads
    Chemistry: Dye-terminator Big Dye; 100% of reads
    Assembly program: Phrap; version 0.960731
    Consensus quality: 183444 bases at least Q40
    Consensus quality: 190666 bases at least Q30
    Consensus quality: 193645 bases at least Q20
    Insert size: 163000; agarose-fp
    Insert size: 195036; sum-of-contigs
    Quality coverage: 9.8 in Q20 bases; agarose-fp
    Quality coverage: 8.2 in Q20 bases; sum-of-contigs
    -----
    * NOTE: This is a 'working draft' sequence. It currently
    * consists of 32 contigs. The true order of the pieces
    * is not known and their order in this sequence record is
    * arbitrary. Gaps between the contigs are represented as
    * runs of N, but the exact sizes of the gaps are unknown.
    * This record will be updated with the finished sequence
    * as soon as it is available and the accession number will
    *
    FEATURES
    source
    1..198136
    Location/Qualifiers
    49624: contig of 49624 bp in length
    43724: gap of 100 bp
    50559: contig of 835 bp in length
    50659: gap of 100 bp
    51365: contig of 706 bp in length
    51465: gap of 100 bp
    52165: contig of 700 bp in length
    52265: gap of 100 bp
    53118: contig of 853 bp in length
    53218: gap of 100 bp
    53967: contig of 749 bp in length
    54067: gap of 100 bp
    55196: contig of 1129 bp in length
    55296: gap of 100 bp
    56063: contig of 767 bp in length
    56163: gap of 100 bp
    56986: contig of 823 bp in length
    57086: gap of 100 bp
    57588: contig of 672 bp in length
    57598: gap of 100 bp
    58532: contig of 674 bp in length
    58632: gap of 100 bp
    59853: contig of 1353 bp in length
    60085: gap of 100 bp
    61366: contig of 1281 bp in length
    61466: gap of 100 bp
    62384: contig of 918 bp in length
    62484: gap of 100 bp
    63865: contig of 1381 bp in length
    63965: gap of 100 bp
    65265: contig of 1300 bp in length
    65365: gap of 100 bp
    66995: contig of 1530 bp in length
    66995: gap of 100 bp
    68299: contig of 1304 bp in length
    68399: gap of 100 bp
    69389: contig of 990 bp in length
    69489: gap of 100 bp
    70667: contig of 1178 bp in length
    70767: gap of 100 bp
    72600: contig of 1833 bp in length
    72700: gap of 100 bp
    73701: contig of 1001 bp in length
    73801: gap of 100 bp
    74876: contig of 1075 bp in length
    74976: gap of 100 bp
    76790: contig of 1814 bp in length
    76890: gap of 100 bp
    78031: contig of 1141 bp in length
    78131: gap of 100 bp
    79413: contig of 1282 bp in length
    83211: gap of 100 bp
    83212: contig of 3698 bp in length
    83311: gap of 100 bp
    86200: contig of 2889 bp in length
    86300: gap of 100 bp
    93593: contig of 7233 bp in length
    93693: gap of 100 bp
    115200: contig of 21507 bp in length
    115201: contig of 100 bp
    172420: contig of 57120 bp in length
    172520: gap of 100 bp
    198136: contig of 25616 bp in length.
    
```


[illegible]

human genes
Unpublished (2004)

Contact: Spindel ER
Division of Neuroscience
Oregon National Primate Research Center
505 NW 185th Avenue, Beaverton, OR 97006, USA
Tel: 403-690-5388
Fax: 503-690-5384
Email: spindel@ohsu.edu
Primer A: atactgtgaggaattgact
Primer B: attacttgggaagttaaaaa
STS size: 458

PCR Profile:
Hot Start: 95 degrees C for 2.00 min
Denaturation: 95 degrees C for 0.50 min
Annealing: 51 degrees C for 0.50 min
Polymerization: 72 degrees C for 1.00 min
PCR Cycles: 35
Extension: 72 degrees C for 7.0 min
Thermal Cycler: MJ Instruments PTC100

Protocol:
Template: 200 ng
Primer: each 1uM
dNTP's: each 200 uM
Taq Polymerase: 0.05 units/ul (Fast Start High Fidelity, Roche)
Total Vol: 50 ul

Buffer:
MgCl2: 1.8 mM
Fast Start polymerase reaction buffer (Roche)

Bases 5-388 are 96% homologous (Blast) to bases 2198-2587 of NM 002094.1. Primers were chosen to amplify genomic DNA in the 3' region of GSPT1. As human sequence was used to design the primers, the primer sequences are not included in the rhesus sequence provided below. To obtain additional information regarding primers or clones contact: Dr. Robert Norgren; Dept of Genetics, Cell Biology & Anatomy; University of Nebraska Medical Center; 986395 Nebraska Medical Center; Omaha, NE 68198. Email: rnorgren@unmc.edu

A database containing sequences associated with this project can be found at: <http://rhesusgeneschip.unomaha.edu/index.html>.

FEATURES
source
1. .458
/organism="Macaca mulatta"
/mol_type="genomic DNA"
/strain="Indian origin"
/db_xref="taxon:9544"
/clone="MMA6672"
/clone_lib="Rhesus macaque genomic DNA"
/dev_stage="Adult"
/note="Organ: Liver; Vector: pGEM-T Easy; V-type: Plasmid; STS was amplified from rhesus genomic DNA with the human forward and reverse primers listed above and subcloned into pGEM-T Easy"
1. .458
/gene="GSPT1"
/note="G1 to S phase transition 1"
<1. .458
/gene="GSPT1"

gene
STS
1. .458

ORIGIN
Query Match 76.8%; Score 19.2; DB 10; Length 458;
Best Local Similarity 87.5%; Pred. No. 1.5e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AAAAAAATCCTGTAGGTTTCATT 24
|||||
Db 266 AAAAAAATCCTGTAGGTTTCATT 243
|||||

RESULT 14
LOCUS CQ923410/c 545 bp DNA linear PAT 23-NOV-2004
DEFINITION Sequence 4610 from Patent WO2004097052.
ACCESSION CQ923410
VERSION CQ923410.1 GI:56213351
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
AUTHORS Burczynski, M.E., Twine, N.C., Slonim, D.K., Trepicchio, W.L., Strahs, A., Immerman, F. and Dörner, A.J.
TITLE Methods for prognosis and treatment of solid tumors
JOURNAL Patent: WO 2004097052-A 4610 11-NOV-2004;
Wyeth (US); Burczynski, Michael E. (US)
FEATURES
Location/Qualifiers
1. .545
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
64
misc_feature /note="n includes a, c, g, or t, or contains no nucleotide"
69
misc_feature /note="n includes a, c, g, or t, or contains no nucleotide"
83
misc_feature /note="n includes a, c, g, or t, or contains no nucleotide"
354. .378
misc_feature /note="n includes a, c, g, or t, or contains no nucleotide"

ORIGIN
Query Match 76.8%; Score 19.2; DB 6; Length 545;
Best Local Similarity 87.5%; Pred. No. 1.5e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AAAAAAATCCTGTAGGTTTCATT 24
|||||
Db 452 AAAAAAATCCTGTAGGTTTCATT 429
|||||

RESULT 15
LOCUS CQ225863/c 562 bp DNA linear PAT 21-JAN-2004
DEFINITION Sequence 12702 from Patent WO0157273.
ACCESSION CQ225863
VERSION CQ225863.1 GI:41208853
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
AUTHORS Penn, S.G., Hanzel, D.K., Chen, W. and Rank, D.R.
TITLE HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR ANALYSIS OF GENE EXPRESSION IN HUMAN ADULT LIVER<130> PB 0004 WO 3<150> US 60/180,312<151> 04 February 2000 (04.02.00)<150> US 60/207,456<151> 26 May 2000 (26.05.00)<150> US 09/632,366<151> 03 August 2000 (03.08.00)<150> GB 24263.6<151> 27 September 2000 (03.10.00)<150> US 60/236,359<151> 27 September 2000 (27.09.00)<150> US 60/234,687<151> 21 September 2000 (21.09.00)<150> US 09/608,408<151> 30 June 2000 (30.06.00)<170>
Molecular Dynamics Sequence Listing Engine
Patent: WO 0157273-A 12702 09-AUG-2001;
Aeomica, Inc. (US)
FEATURES
Location/Qualifiers
1. .562
source

/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="MAP TO AP000356.1
EXPRESSED IN ADULT LIVER, SIGNAL = 1.7"

ORIGIN

Query Match 76.8%; Score 19.2; DB 6; Length 562;
Best Local Similarity 87.5%; Pred. No. 1.5e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAATCCTGTAGGCTTTTCAGTG 25
|||||
Db 29 AAAAATCCTGTATGCTTTGATTG 6
|||||

RESULT 16

CQ300992/c
LOCUS CQ300992 562 bp DNA linear PAT 23-JAN-2004
DEFINITION Sequence 12097 from Patent WO0186003.
ACCESSION CQ300992
VERSION CQ300992.1 GI:41261569
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 Penn, S.G., Hanzel, D.K., Chen, W. and Rank, D.R.
AUTHORS Human genome-derived single exon nucleic acid probes useful for
analysis of gene expression in human lung
TITLE Patent: WO 0186003-A 12097 15-NOV-2001;
JOURNAL Aeomica, Inc. (US)
FEATURES
Location/Qualifiers
1..562
Source
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="MAP TO AP000356.1
EXPRESSED IN LUNG, SIGNAL = 1.2"

ORIGIN

Query Match 76.8%; Score 19.2; DB 6; Length 562;
Best Local Similarity 87.5%; Pred. No. 1.5e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAATCCTGTAGGCTTTTCAGTG 25
|||||
Db 29 AAAAATCCTGTATGCTTTGATTG 6
|||||

RESULT 17

CQ338275/c
LOCUS CQ338275 562 bp DNA linear PAT 23-JAN-2004
DEFINITION Sequence 12369 from Patent WO0157275.
ACCESSION CQ338275
VERSION CQ338275.1 GI:41287346
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 Penn, S.G., Hanzel, D.K., Chen, W. and Rank, D.R.
AUTHORS Human genome-derived single exon nucleic acid probes useful for
analysis of gene expression in human brain
TITLE Patent: WO 0157275-A 12369 09-AUG-2001;
JOURNAL Aeomica, Inc. (US)
FEATURES
Location/Qualifiers
1..562
Source
/organism="Homo sapiens"

/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="MAP TO AP000356.1
EXPRESSED IN BRAIN, SIGNAL = 5.3"

ORIGIN

Query Match 76.8%; Score 19.2; DB 6; Length 562;
Best Local Similarity 87.5%; Pred. No. 1.5e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAATCCTGTAGGCTTTTCAGTG 25
|||||
Db 29 AAAAATCCTGTATGCTTTGATTG 6
|||||

RESULT 18

BV565426
LOCUS BV565426 711 bp DNA linear STS 09-APR-2005
DEFINITION rc167d10.b1 Clint Pan troglodytes versus STS genomic, sequence
tagged site.
ACCESSION BV565426
VERSION BV565426.1 GI:62456447
KEYWORDS STS.
SOURCE Pan troglodytes versus
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Pan.
1 (bases 1 to 711)
REFERENCE Mikkelson, T.S., Hillier, W.L., Eichler, E.E., Zody, M.C. and
AUTHORS Jaffe, D.B.
TITLE Initial Sequence of the Chimpanzee Genome and Comparison with the
JOURNAL Human Genome
COMMENT Unpublished (2005)
Contact: Michael C. Zody
Broad Institute of MIT and Harvard
320 Charles Street, Cambridge, MA 02141, USA
Tel: 6172580933
Fax: 6172580903
Email: mczody@broad.mit.edu
Primer A: No sequence submitted
Primer B: No sequence submitted
STS size: 711
Protocol:
23,021,928 chimpanzee whole genome shotgun reads were aligned to
the Human genome NCBI
Build 34 (hg16, July 2003). Chimp WGS reads were from 9 donors,
including Clint (Pan
troglodytes versus), 3 other Pan troglodytes versus chimps
(Donald, Karlien, Yvonne), 3 pan
troglodytes troglodytes chimps (Noemie, Masuku, Clara) and 2 chimps
of unknown origin
(Gon, Unknown Chimp). Common names: Pan troglodytes versus is the
western chimp and Pan
troglodytes troglodytes is the central chimp. To be included in
chimpanzee SNP discovery, a
read must be at least 500bp in length, at least 50% of its base
calls must have Phred
score >= 20, at least 30% of its base calls must satisfy
SNQS(30,25) (single strand NQS, the
base in question has Phred score >= 30, the surrounding 10 bases in
the read have Phred
score >= 25), and the read must have at least 200 bp SNQS(30,25)
bases. Reads not uniquely
placed in the genome and read pairs whose two ends were not
consistently placed were
discarded. After above filtering, NQS(30,25) standard was applied
to all pairs of
overlapping reads to call NQS bases and SNPs. Alignments (between
two reads) with less
than 100 NQS bases or with SNP rate > 0.01 were discarded. To
exclude alignment between two

copies of a single read, comparisons between two reads that share 95% of their genome alignments (>=95% bases of read A and >=95% bases of read B were placed at the same locus of human genome) were discarded.

FEATURES
source
1. .711
/organism="Pan troglodytes verus"
/mol_type="genomic DNA"
/sub_species="verus"
/db_xref="taxon:37012"
/clone_lib="Clint"
<1. .>711

STS
ORIGIN
Query Match 76.8%; Score 19.2; DB 10; Length 711;
Best Local Similarity 87.5%; Pred. No. 1.5e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 2 AAAAACTCCTGTAGGCTTTCAGTG 25
Db 315 AAAAACTCCTGTATGCTTTGATTG 338

RESULT 19
BV592074 785 bp DNA linear STS 12-APR-2005
LOCUS G591P51714RD5.T0 Clint Pan troglodytes verus STS genomic, sequence tagged site.
DEFINITION BV592074
ACCESSION BV592074.1 GI:62507803
VERSION
KEYWORDS
SOURCE
ORGANISM
Pan troglodytes verus
Pan troglodytes verus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Pan.
REFERENCE 1 (bases 1 to 785)
AUTHORS Mikkelsen,T.S., Hillier,W.L., Eichler,E.E., Zody,M.C. and Jaffe,D.B.
TITLE Initial Sequence of the Chimpanzee Genome and Comparison with the Human Genome
JOURNAL Unpublished (2005)
COMMENT

Contact: Michael C. Zody
Broad Institute of MIT and Harvard
320 Charles Street, Cambridge, MA 02141, USA
Tel: 6172580933
Fax: 6172580903
Email: mczody@broad.mit.edu
Primer A: No sequence submitted
Primer B: No sequence submitted
STS size: 785
Protocol:
23, 021,928 chimpanzee whole genome shotgun reads were aligned to the Human genome NCBI
Build 34 (hg16,July 2003). Chimp WGS reads were from 9 donors, including Clint (Pan troglodytes verus), 3 other Pan troglodytes verus chimps (Donald,Karlien,Yvonne), 3 Pan troglodytes troglodytes chimps (Noemie,Masuku,Clara) and 2 chimps of unknown origin
(Gen,Unknown Chimp). Common names: Pan troglodytes verus is the western chimp and Pan troglodytes troglodytes is the central chimp. To be included in chimpanzee SNP discovery, a read must be at least 500bp in length, at least 50% of its base calls must have Phred score >= 20, at least 30% of its base calls must satisfy SNQS(30,25) (single strand NQS, the base in question has Phred score >= 30, the surrounding 10 bases in the read have Phred score >= 25), and the read must have at least 200 bp SNQS(30,25)

bases. Reads not uniquely placed in the genome and read pairs whose two ends were not consistently placed were discarded. After above filtering, NQS(30,25) standard was applied to all pairs of overlapping reads to call NQS bases and SNPs. Alignments (between two reads) with less than 100 NQS bases or with SNP rate > 0.01 were discarded. To exclude alignment between two copies of a single read, comparisons between two reads that share 95% of their genome alignments (>=95% bases of read A and >=95% bases of read B were placed at the same locus of human genome) were discarded.

FEATURES
Location/Qualifiers
1. .785
/organism="Pan troglodytes verus"
/mol_type="genomic DNA"
/sub_species="verus"
/db_xref="taxon:37012"
/clone_lib="Clint"
<1. .>785

STS
ORIGIN
Query Match 76.8%; Score 19.2; DB 10; Length 785;
Best Local Similarity 87.5%; Pred. No. 1.5e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 2 AAAAACTCCTGTAGGCTTTCAGTG 25
Db 184 AAAAACTCCTGTATGCTTTGATTG 207

RESULT 20
AB003502/c 2079 bp mRNA linear ROD 27-AUG-1998
LOCUS Mus musculus mRNA for Guanine Nucleotide Regulatory Protein, partial cds.
DEFINITION AB003502
ACCESSION AB003502.1 GI:3461879
VERSION
KEYWORDS Guanine Nucleotide Regulatory Protein.
SOURCE Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 2079)
AUTHORS Hoshino,S., Imai,M., Mizutani,M., Kikuchi,Y., Hanaoka,F., Ui,M. and Katada,T.
TITLE Molecular cloning of a novel member of the eukaryotic polypeptide chain-releasing factors (eRF). Its identification as eRF3 interacting with eRF1
JOURNAL J. Biol. Chem. 273 (35), 22254-22259 (1998)
PUBMED 9712840
REFERENCE 2 (bases 1 to 2079)
AUTHORS Hoshino,S.
TITLE Direct Submission
JOURNAL Submitted (02-MAY-1997) Shin-ichi Hoshino, Graduate School of Pharmaceutical Sciences University of Tokyo, Physiological Chemistry; 7-3-1 Hongo, Bunkyo-ku, Tokyo 113, Japan
(E-mail:hoshino@mol.f.u.-tokyo.ac.jp, Tel:03-3812-2111, Fax:03-3815-9604)

FEATURES
Location/Qualifiers
1. .2079
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
<1. .>1653
/codon_start=1
/product="Guanine Nucleotide Regulatory Protein"
/protein_id="BAAJ32526.1"
/db_xref="GI:3461880"
/translation="VPSFLRGPAPPLSPAGAGDHDGAGGAGGPEVPVSSQDQSC

Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAL Plate: 2 Row: h Column: 11

This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4504166.

```
FEATURES             source
  Location/Qualifiers
    1..2523
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="MGC:1735 IMAGE:2822947"
      /tissue_type="lung, small cell carcinoma"
      /clone_lib="NIH_MGC_7"
      /lab_host="DH10B-R"
      /note="Vector: pOTB7"
    1..2523
      /gene="GSTP1"
      /gene="synonyms: GST1, ETF3A, eRF3a"
      /db_xref="GeneID:2935"
      /db_xref="MIM:139259"
    159..2060
      /gene="GSTP1"
      /codon_start=1
      /product="GSTP1 protein"
      /protein_id="AAH09503.2"
      /db_xref="GI:33874734"
      /db_xref="GeneID:2935"
      /db_xref="MIM:139259"
      /translation="WDPSGGGGGGSSGSSSDSAPDCWDQADMEAPGPGCGGG
      GSLAARAAQRNLAAPSRQLNVNAKPPVNVHAAEFVPSFLRCPAAPPPAGGAAN
      NHGAGSAGGAAAPVSSQEOESLCEGNSAVSMELSPTENGETEMSPESWEHKEE
      ISEAPFGGSLGDPPEESAHEMMEEEIEIPKPSVAPGAPKKEHVNVVFIQHVD
      AGKSTIGGIMVLTGVKDKRTLEKEREAKENRTWYLSWALDTNQBERDKGTVEV
      GRAYPETAKHTFTILDAPGHKSFVNPMIGGASQADLAVLVSARGEFETGFKGGQT
      RPHAMTAGVKHLVLILNKMDPTVNMNSRYECKELVPFLKKVGFNPKKDHF
      MPCSILTGANLKEQDQFCWYIGLFFIPLYDLNLPNRSVDGPIRLPIDVKYKDMGTV
      VLGKLESSICGQQLVMPNKHNVELGILSDDETDTVAPEGNLKRLKGIEBEEI
      LPFGFLCPNNLCGSRGTFDAQIVIEHKHSIICPGYNVAVLHIHTCIEEVEITALLCLV
      DKKSGEKSKTRPRFVKQDQVCILARTAGTICLETFFKFPQPMGRFTLRDEGKTIAGK
      VLKLVPEKD"
    774..2048
      /gene="GSTP1"
      /note="TEFI; Region: Translation elongation factor
      EF-1alpha (GTPase) [Translation, ribosomal structure and
      biogenesis]"
      /db_xref="CDD:COG5256"
```

```
ORIGIN
Query Match          76.8%; Score 19.2; DB 8; Length 2523;
Best Local Similarity 87.5%; Pred. No. 1.5e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAACTCCTGTAGCTTTCAGT 24
      ||||| ||||| ||||| ||||| |||||
Db 2375 AAAAAATTCCTGTAGTTTTCATT 2352
```

```
RESULT 24
AX833607/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
```

```
REFERENCE
1 Isogai,T., Sugiyama,T., Otsuki,T., Wakamatsu,A., Sato,H., Ishii,S.,
AUTHORS
```

```
Yamamoto,J.I., Isono,Y., Hio,Y., Otsuka,K., Nagai,K., Irie,R.,
Tanechika,I., Seki,N., Yoshikawa,T., Otsuka,M., Nagahari,K. and
Masuo,Y.
Full-length cDNA sequences
Patent: EP 1347046-A 731 24-SEP-2003;
Research Association for Biotechnology (JP)
Location/Qualifiers
  1..2527
    /organism="Homo sapiens"
    /mol_type="unassigned DNA"
    /db_xref="taxon:9606"

ORIGIN
Query Match          76.8%; Score 19.2; DB 6; Length 2527;
Best Local Similarity 87.5%; Pred. No. 1.5e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAACTCCTGTAGCTTTCAGT 24
      ||||| ||||| ||||| ||||| |||||
Db 2405 AAAAAATTCCTGTAGTTTTCATT 2382

RESULT 25
AK095483/c
LOCUS
DEFINITION
Homo sapiens cDNA FLJ38164 fis, clone DFNES2003742, highly similar
to G1 TO S PHASE TRANSITION PROTEIN 1.
ACCESSION
VERSION
KEYWORDS
SOURCE
Homo sapiens
Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE
1 Ota,T., Suzuki,Y., Nishikawa,T., Otsuki,T., Sugiyama,T., Irie,R.,
Wakamatsu,A., Hayashi,K., Sato,H., Nagai,K., Kimura,K., Makita,H.,
Sekine,M., Oyabashi,M., Nishi,T., Shibahara,T., Tanaka,T.,
Ishii,S., Yamamoto,J., Saiko,K., Kawai,Y., Isono,Y., Nakamura,Y.,
Nagahari,K., Murakami,K., Yasuda,T., Iwayanagi,T., Wagatsuma,M.,
Shiratori,A., Sudo,H., Hosoiri,T., Kaku,Y., Kodaira,H., Kondo,H.,
Sugawara,M., Takahashi,M., Kanda,K., Yokoi,T., Furuya,T.,
Kikkawa,E., Omura,Y., Abe,K., Kamihara,K., Katsuta,N., Sato,K.,
Tanikawa,M., Yamazaki,M., Ninomiya,K., Iehibashi,T., Yamashita,H.,
Murakawa,K., Fujimori,K., Tanai,H., Kimata,M., Watanabe,M.,
Hiraoka,S., Chiba,Y., Ishida,S., Ono,Y., Takiguchi,S., Watanabe,S.,
Yosida,M., Hotuta,T., Kusano,J., Kanehori,K., Takahashi-Fujii,A.,
Hara,H., Tanase,T., Nomura,Y., Togiya,S., Komai,F., Hara,R.,
Takeuchi,K., Arita,M., Imose,N., Musashino,K., Yuuki,H., Oshima,A.,
Sasaki,N., Aotsuka,S., Yoshikawa,Y., Matsunawa,H., Ichihara,T.,
Shiohata,N., Sano,S., Moriya,S., Momiya,H., Satoh,N., Takami,S.,
Terashima,Y., Suzuki,O., Nakagawa,S., Senoh,A., Mizoguchi,H.,
Goto,Y., Shimizu,F., Wakebe,H., Hishigaki,H., Watanabe,T.,
Sugiyama,A., Takemoto,M., Kawakami,B., Yamazaki,M., Watanabe,K.,
Kumagai,A., Itakura,S., Fukuzumi,Y., Fujimori,Y., Komiyama,M.,
Tashiro,H., Tanigami,A., Fujiwara,T., Ono,T., Yamada,K., Fujii,Y.,
Ozaki,K., Hirao,M., Ohmori,Y., Kawabata,A., Hikiji,T., Kobatake,N.,
Inagaki,H., Ikema,Y., Okamoto,S., Okitani,R., Kawakami,T.,
Noguchi,S., Itoh,T., Shigeta,K., Senba,T., Matsumura,K.,
Nakajima,Y., Mizuno,T., Morinaga,M., Sasaki,M., Togashi,T.,
Oyama,M., Hata,H., Watanabe,M., Komatsu,T., Mizushima-Sugano,J.,
Satoh,T., Shirai,Y., Takahashi,Y., Nakagawa,K., Okumura,K.,
Nagase,T., Nomura,N., Kikuchi,H., Masuno,Y., Yamashita,R.,
Nakai,K., Yada,T., Nakamura,Y., Ohara,O., Isogai,T. and Sugano,S.
Complete sequencing and characterization of 21,243 full-length
human cDNAs
Nat. Genet. 36 (1), 40-45 (2004)

JOURNAL
PUBMED
REFERENCE
1 Suzuki,O., Sasaki,N., Aotsuka,S., Shoji,T., Ichihara,T.,
Shiohata,N., Matsumoto,K., Hirano,M., Sano,S., Nomura,R.,
Yoshikawa,Y., Matsumura,Y., Moriya,S., Chiba,E., Momiya,H.,
```

Onogawa,S., Kaeriyama,S., Satoh,N., Matsunawa,H., Takahashi,E., Katsoka,R., Kuga,N., Kuroda,A., Satoh,I., Kamata,K., Takami,S., Terashima,Y., Watanabe,M., Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H., Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K., Wagatsuma,M., Murakawa,K., Kaneshiro,K., Takahashi-Fujii,A., Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y., Sugano,S., Nagahari,K., Masuho,Y., Nagai,K. and Isogai,T.
NEDO human cDNA sequencing project
Unpublished
3 (bases 1 to 2527)
Isogai,T. and Yamamoto,J.
Direct Submission
Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
(E-mail:genomics@ri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.
FEATURES
source
LOCATION/Qualifiers
1..2527
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DFNES2003742"
/cell_type="normal dermal fibroblasts (Neonatal Skin) (NHDF2564)"
/clone_lib="DFNES2"
/notes="cloning vector: pME18SFL3
primary culture, normal dermal fibroblasts"
ORIGIN
Query Match 76.8%; Score 19.2; DB 8; Length 2527;
Best Local Similarity 87.5%; Pred. No. 1.5e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 AAAAACTCCTGTAGGCTTTTCAGT 24
||||| ||||||| ||||||| ||||||| |||||||
Db 2405 AAAAAATTCCTGTAGGTTTCATT 2382
||||| ||||||| ||||||| ||||||| |||||||
RESULT 26
CQ714266/c 2586 bp DNA linear PAT 03-FEB-2004
LOCUS
DEFINITION
Sequence 200 from Patent WO2068579.
ACCESSION
CQ714266
VERSION
CQ714266.1 GI:42275123
KEYWORDS
Homo sapiens (human)
SOURCE
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1
Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
Kits, such as nucleic acid arrays, comprising a majority of humanexons or transcripts, for detecting expression and other uses thereof
JOURNAL
Patent: WO 02068579-A 200 06-SEP-2002;
PE Corporation (NY) (US)
FEATURES
source
LOCATION/Qualifiers
1..2586
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 76.8%; Score 19.2; DB 6; Length 2586;
Best Local Similarity 87.5%; Pred. No. 1.5e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 AAAAACTCCTGTAGGCTTTTCAGT 24
||||| ||||||| ||||||| ||||||| |||||||
Db 2405 AAAAAATTCCTGTAGGTTTCATT 2382
||||| ||||||| ||||||| ||||||| |||||||
RESULT 26
CQ714266/c 2586 bp DNA linear PAT 10-MAR-2005
LOCUS
DEFINITION
Sequence 486 from Patent WO2005016962.
ACCESSION
CS030980
VERSION
CS030980.1 GI:60731250
KEYWORDS
Homo sapiens (human)
SOURCE
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1
Abbas,A., Clark,H., Ouyang,W., Williams,M.P., Wood,W.I. and Wu,T.D.
Compositions and methods for the treatment of immune related diseases
JOURNAL
Patent: WO 2005016962-A 486 24-FEB-2005;
Genentech, Inc. (US)
FEATURES
source
LOCATION/Qualifiers
1..2587
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 76.8%; Score 19.2; DB 6; Length 2587;
Best Local Similarity 87.5%; Pred. No. 1.5e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 AAAAACTCCTGTAGGCTTTTCAGT 24
||||| ||||||| ||||||| ||||||| |||||||
Db 2463 AAAAAATTCCTGTAGGTTTCATT 2440
||||| ||||||| ||||||| ||||||| |||||||
RESULT 28
CS030980/c 2587 bp DNA linear PAT 10-MAR-2005
LOCUS
DEFINITION
Sequence 486 from Patent WO2005016962.
ACCESSION
CS030980
VERSION
CS030980.1 GI:60731250
KEYWORDS
Homo sapiens (human)
SOURCE
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1
Abbas,A., Clark,H., Ouyang,W., Williams,M.P., Wood,W.I. and Wu,T.D.
Compositions and methods for the treatment of immune related diseases
JOURNAL
Patent: WO 2005016962-A 486 24-FEB-2005;
Genentech, Inc. (US)
FEATURES
source
LOCATION/Qualifiers
1..2587
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 76.8%; Score 19.2; DB 6; Length 2587;
Best Local Similarity 87.5%; Pred. No. 1.5e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 AAAAACTCCTGTAGGCTTTTCAGT 24
||||| ||||||| ||||||| ||||||| |||||||
Db 2463 AAAAAATTCCTGTAGGTTTCATT 2440
||||| ||||||| ||||||| ||||||| |||||||

Query Match 76.8%; Score 19.2; DB 6; Length 2586;
Best Local Similarity 87.5%; Pred. No. 1.5e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 AAAAACTCCTGTAGGCTTTTCAGT 24
||||| ||||||| ||||||| ||||||| |||||||
Db 2464 AAAAAATTCCTGTAGGTTTCATT 2441
||||| ||||||| ||||||| ||||||| |||||||
RESULT 27
CQ981521/c 2587 bp DNA linear PAT 25-JAN-2005
LOCUS
DEFINITION
Sequence 376 from Patent EP1498424.
ACCESSION
CQ981521
VERSION
CQ981521.1 GI:58190811
KEYWORDS
Homo sapiens (human)
SOURCE
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1
Rosenthal,A., Hermann,K., Heiden,E., Pilarsky,C., Bruemendorf,T., Staub,E., Roepcke,S., Mennerich,D., Kinnemann,H. and Li,X.
Human nucleic acid sequences from lung tumours
Patent: EP 1498424-A 376 19-JAN-2005;
Hinzmann, Bernd (DE); Hermann, Klaus (DE); Heiden, Esmeralda (DE); Rosenthal, Andre (DE)
FEATURES
source
LOCATION/Qualifiers
1..2587
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 76.8%; Score 19.2; DB 6; Length 2587;
Best Local Similarity 87.5%; Pred. No. 1.5e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 AAAAACTCCTGTAGGCTTTTCAGT 24
||||| ||||||| ||||||| ||||||| |||||||
Db 2463 AAAAAATTCCTGTAGGTTTCATT 2440
||||| ||||||| ||||||| ||||||| |||||||
RESULT 28
CS030980/c 2587 bp DNA linear PAT 10-MAR-2005
LOCUS
DEFINITION
Sequence 486 from Patent WO2005016962.
ACCESSION
CS030980
VERSION
CS030980.1 GI:60731250
KEYWORDS
Homo sapiens (human)
SOURCE
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1
Abbas,A., Clark,H., Ouyang,W., Williams,M.P., Wood,W.I. and Wu,T.D.
Compositions and methods for the treatment of immune related diseases
JOURNAL
Patent: WO 2005016962-A 486 24-FEB-2005;
Genentech, Inc. (US)
FEATURES
source
LOCATION/Qualifiers
1..2587
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 76.8%; Score 19.2; DB 6; Length 2587;
Best Local Similarity 87.5%; Pred. No. 1.5e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 AAAAACTCCTGTAGGCTTTTCAGT 24
||||| ||||||| ||||||| ||||||| |||||||
Db 2463 AAAAAATTCCTGTAGGTTTCATT 2440
||||| ||||||| ||||||| ||||||| |||||||

```
QY 1 AAAAACTCCTGTAGGCTTTTCAGT 24
|||||
Db 2463 AAAAAATTCCTGTAGGTTTTCATT 2440

RESULT 29
CS036801/c
LOCUS 2587 bp DNA linear PAT 10-MAR-2005
DEFINITION Sequence 6307 from Patent WO2005016962.
ACCESSION CS036801
VERSION CS036801.1 GI:60734216
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
AUTHORS Abbas,A., Clark,H., Ouyang,W., Williams,M.P., Wood,W.I. and Wu,T.D.
TITLE Compositions and methods for the treatment of immune related
diseases
JOURNAL Patent: WO 2005016962-A 6307 24-FEB-2005;
Genentech, Inc. (US)
FEATURES
source
1..2587
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 76.8%; Score 19.2; DB 6; Length 2587;
Best Local Similarity 87.5%; Pred. No. 1.5e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAACTCCTGTAGGCTTTTCAGT 24
|||||
Db 2463 AAAAAATTCCTGTAGGTTTTCATT 2440

RESULT 30
CS039932/c
LOCUS 2587 bp DNA linear PAT 22-MAR-2005
DEFINITION Sequence 486 from Patent WO2005019258.
ACCESSION CS039932
VERSION CS039932.1 GI:61847833
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
AUTHORS Abbas,A., Clark,H., Ouyang,W., Williams,P.M., Wood,W.I. and Wu,T.D.
TITLE Compositions and methods for the treatment of immune related
diseases
JOURNAL Patent: WO 2005019258-A 486 03-MAR-2005;
Genentech, Inc. (US)
FEATURES
source
1..2587
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 76.8%; Score 19.2; DB 6; Length 2587;
Best Local Similarity 87.5%; Pred. No. 1.5e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAACTCCTGTAGGCTTTTCAGT 24
|||||
Db 2463 AAAAAATTCCTGTAGGTTTTCATT 2440

RESULT 31
CS045753/c
LOCUS 2587 bp DNA linear PAT 22-MAR-2005
DEFINITION Sequence 6307 from Patent WO2005019258.
ACCESSION CS045753
VERSION CS045753.1 GI:61852019
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
AUTHORS Abbas,A., Clark,H., Ouyang,W., Williams,P.M., Wood,W.I. and Wu,T.D.
TITLE Compositions and methods for the treatment of immune related
diseases
JOURNAL Patent: WO 2005019258-A 6307 03-MAR-2005;
Genentech, Inc. (US)
FEATURES
source
1..2587
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 76.8%; Score 19.2; DB 6; Length 2587;
Best Local Similarity 87.5%; Pred. No. 1.5e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAACTCCTGTAGGCTTTTCAGT 24
|||||
Db 2463 AAAAAATTCCTGTAGGTTTTCATT 2440

RESULT 32
CS086305/c
LOCUS 2587 bp DNA linear PAT 25-MAY-2005
DEFINITION Sequence 8 from Patent WO2005042786.
ACCESSION CS086305
VERSION CS086305.1 GI:66711877
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
AUTHORS Harris,C. and Davis,L.
TITLE Compositions and methods for glioma classification
JOURNAL Patent: WO 2005042786-A 8 12-MAY-2005;
Exagen Diagnostics, Inc. (US)
FEATURES
source
1..2587
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 76.8%; Score 19.2; DB 6; Length 2587;
Best Local Similarity 87.5%; Pred. No. 1.5e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAACTCCTGTAGGCTTTTCAGT 24
|||||
Db 2463 AAAAAATTCCTGTAGGTTTTCATT 2440

RESULT 33
HSGST1/c
LOCUS 2587 bp mRNA linear PRI 18-APR-2005
DEFINITION Human GST1-Hs mRNA for GTP-binding protein.
ACCESSION X17644
VERSION X17644.1 GI:31920
```


Center: RIKEN Genomic Sciences Center (GSC)
Center code: RIKEN
Web site: <http://hgp.gsc.riken.go.jp/>
Contact: hattori@gs.riken.go.jp
----- Project Information -----
Center project name: HumDraft11
Center clone name: XXPac-662A14
----- Summary Statistics -----
Sequencing vector: PCR products; 100% of reads
Chemistry: Dye-terminator ET-amersham; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 103806 bases at least Q40
Consensus quality: 109449 bases at least Q30
Consensus quality: 111869 bases at least Q20
Insert size: 113473; sum-of-contigs
Quality coverage: 4.03x in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of 29 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved

```
1 38138 contig of 38138 bp in length
38239 44698 contig of 4460 bp in length
44799 49514 contig of 4716 bp in length
49615 55534 contig of 5920 bp in length
55635 60136 contig of 4502 bp in length
60237 65623 contig of 5387 bp in length
65724 70102 contig of 4379 bp in length
70203 71387 contig of 1185 bp in length
71488 73839 contig of 2352 bp in length
73940 77366 contig of 3427 bp in length
77467 80065 contig of 2599 bp in length
80166 83825 contig of 3660 bp in length
83926 86777 contig of 1073 bp in length
86878 89883 contig of 100 bp
89984 92751 contig of 2868 bp in length
92852 95068 contig of 2217 bp in length
95169 96739 contig of 100 bp
96840 97912 contig of 1571 bp in length
98013 98744 contig of 100 bp
98845 101153 contig of 1309 bp in length
101254 102868 contig of 1615 bp in length
102969 104921 contig of 100 bp
104922 105021 contig of 1479 bp in length
105022 106500 contig of 100 bp
106501 108529 contig of 1929 bp in length
108530 108630 contig of 2212 bp in length
108631 110841 contig of 1426 bp in length
110842 112367 contig of 100 bp
112368 113665 contig of 1198 bp in length
113666 114974 contig of 1209 bp in length
114975 115075 contig of 1199 bp in length
115076 116273 contig of 1199 bp in length
Sequence updated (26-May-2000).
```

* NOTE: This is a 'working draft' sequence. It currently consists of 29 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

```
1 38138 contig of 38138 bp in length
38239 44698 contig of 4460 bp in length
44799 49514 contig of 4716 bp in length
49615 55534 contig of 5920 bp in length
55635 60136 contig of 4502 bp in length
60237 65623 contig of 5387 bp in length
65724 70102 contig of 4379 bp in length
70203 71387 contig of 1185 bp in length
71488 73839 contig of 2352 bp in length
73940 77366 contig of 3427 bp in length
77467 80065 contig of 2599 bp in length
80166 83825 contig of 3660 bp in length
83926 86777 contig of 1073 bp in length
86878 89883 contig of 100 bp
89984 92751 contig of 2868 bp in length
92852 95068 contig of 2217 bp in length
95169 96739 contig of 100 bp
96840 97912 contig of 1571 bp in length
98013 98744 contig of 100 bp
98845 101153 contig of 1309 bp in length
101254 102868 contig of 1615 bp in length
102969 104921 contig of 100 bp
104922 105021 contig of 1479 bp in length
105022 106500 contig of 100 bp
106501 108529 contig of 1929 bp in length
108530 108630 contig of 2212 bp in length
108631 110841 contig of 1426 bp in length
110842 112367 contig of 100 bp
112368 113665 contig of 1198 bp in length
113666 114974 contig of 1209 bp in length
114975 115075 contig of 1199 bp in length
Sequence updated (26-May-2000).
```

```
* 65624 65723: gap of 100 bp
* 65724 70102: contig of 4379 bp in length
* 70103 70202: gap of 100 bp
* 70203 71387: contig of 1185 bp in length
* 71388 71487: gap of 100 bp
* 71488 73839: contig of 2352 bp in length
* 73840 73939: gap of 100 bp
* 73940 77366: contig of 3427 bp in length
* 77367 77466: gap of 100 bp
* 77467 80065: contig of 2599 bp in length
* 80066 80165: gap of 100 bp
* 80166 83825: contig of 3660 bp in length
* 83826 83925: gap of 100 bp
* 83926 86777: contig of 2852 bp in length
* 86778 86777: gap of 100 bp
* 86878 89783: contig of 2306 bp in length
* 89784 89883: gap of 100 bp
* 89884 92751: contig of 2868 bp in length
* 92752 92851: gap of 100 bp
* 92852 95068: contig of 2217 bp in length
* 95069 95168: gap of 100 bp
* 95169 96739: contig of 1571 bp in length
* 96740 96839: gap of 100 bp
* 96840 97912: contig of 1073 bp in length
* 97913 98012: gap of 100 bp
* 98013 98744: contig of 1732 bp in length
* 98745 99844: gap of 100 bp
* 99845 101153: contig of 1309 bp in length
* 101154 101253: gap of 100 bp
* 101254 102868: contig of 1615 bp in length
* 102869 102968: gap of 100 bp
* 102969 104921: contig of 1953 bp in length
* 104922 105021: gap of 100 bp
* 105022 106500: contig of 1479 bp in length
* 106501 106600: gap of 100 bp
* 106601 108529: contig of 1929 bp in length
* 108530 108630: gap of 100 bp
* 108631 110841: contig of 2212 bp in length
* 110842 110941: gap of 100 bp
* 110942 112367: contig of 1426 bp in length
* 112368 112467: gap of 100 bp
* 112468 113665: contig of 1198 bp in length
* 113666 113765: gap of 100 bp
* 113766 114974: contig of 1209 bp in length
* 114975 115074: gap of 100 bp
* 115075 116273: contig of 1199 bp in length.
FEATURES
    source
        1..116273
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            /chromosome="11"
            /map="11q13"
            /clone="XXPac-662A14"
        1..38138
            /note="assembly_fragment"
        38239..44698
            /note="assembly_fragment"
        44799..49514
            /note="assembly_fragment"
        49615..55534
            /note="assembly_fragment"
        55635..60136
            /note="assembly_fragment"
        60237..65623
            /note="assembly_fragment"
        65724..70102
            /note="assembly_fragment"
        70203..71387
            /note="assembly_fragment"
        71488..73839
            /note="assembly_fragment"
        73940..77366
            /note="assembly_fragment"
```

```

misc_feature /note="assembly_fragment"
77467..80065
misc_feature /note="assembly_fragment"
80166..83825
misc_feature /note="assembly_fragment"
83926..86777
misc_feature /note="assembly_fragment"
86878..89783
misc_feature /note="assembly_fragment"
89884..92751
misc_feature /note="assembly_fragment"
92852..95068
misc_feature /note="assembly_fragment"
95169..96739
misc_feature /note="assembly_fragment"
96840..97912
misc_feature /note="assembly_fragment"
98013..99744
misc_feature /note="assembly_fragment"
99845..101153
misc_feature /note="assembly_fragment"
101254..102868
misc_feature /note="assembly_fragment"
102969..104921
misc_feature /note="assembly_fragment"
105022..106500

Query Match 76.8%; Score 19.2; DB 14; Length 116273;
Best Local Similarity 87.5%; Pred. No. 1.5e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAATCTCTGTTAGGCTTCAGTG 25
DB 99401 AAAAATCTCAGGAGGCTCAGTG 99378

RESULT 40
HSU95742/c
LOCUS HSU95742 139376 bp DNA linear PRI 10-JAN-2000
DEFINITION Human chromosome 16p13.1 BAC clone CIT987SK-551G9 complete
sequence.
ACCESSION U95742
VERSION U95742.1 GI:2339843
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
REFERENCE 1 (bases 1 to 139376)
AUTHORS Loftus,B.J., Kim,U.J., Sneddon,V.P., Kalush,F., Brandon,R.,
Fuhrmann,J., Mason,T., Crosby,M.L., Barnstead,M., Cronin,L.,
Deslattes Mays,A., Cao,Y., Xu,R.X., Kang,H.L., Mitchell,S.,
Eichler,E.E., Harris,P.C., Venter,J.C. and Adams,M.D.
Genome duplications and other features in 12 Mb of DNA sequence
from human chromosome 16p and 16q
Genomics 60 (3), 295-308 (1999)
JOURNAL 10493829
PUBMED 2 (bases 1 to 139376)
REFERENCE Adams,M.D., Loftus,B.J., Zhou,L., Phillips,C., Brandon,R.,
Fuhrmann,J., Kim,U.J., Kerlavage,A.R. and Venter,J.C.
Human chromosome 16p13.1 BAC clone CIT987SK-551G9 complete sequence
Unpublished
JOURNAL 3 (bases 1 to 139376)
REFERENCE Adams,M.D.
JOURNAL Direct Submission
TITLE Submitted (28-MAR-1997) The Institute for Genomic Research, 9712
Medical Center Dr., Rockville, MD 20850, USA
REFERENCE 4 (bases 1 to 139376)
AUTHORS Adams,M.D., Loftus,B.J., Zhou,L., La Bombard,M., Kim,U.J. and
Venter,J.C.
JOURNAL Direct Submission
TITLE Submitted (21-AUG-1997) The Institute for Genomic Research, 9712
Medical Center Dr., Rockville, MD 20850, USA

COMMENT
Medical Center Dr., Rockville, MD 20850, USA
On Aug 21, 1997 this sequence version replaced gi:1930144.
BAC clone CIT987SK-551G9 is located in band 16p13.1 of chromosome
16. Genes were identified by a combination of five methods: XGRAIL
(available by anonymous ftp from arthur.epm.ornl.gov), Genefinder
(available by anonymous ftp from colin@u.washington.edu), GENSCAN
(searches of the EST database at TIGR
http://www.tigr.org/tdb/hcd/hcd.html) and searches against a
peptide database. Repeats were identified using Censor (Jurka, J.,
Klonowski, P., Dagman, V., Pelton, P. Censor-a program for the
identification and elimination of repetitive elements from DNA
sequences. Computers Chem 20: 119-121 (1996); available by
anonymous ftp from ncbi.nlm.nih.gov).
FEATURES
Location/Qualifiers
1..139376
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="16"
/map="16p13.1"
/clone="551G9"
complement(849..1113)
/rpt_family="Alu-Sz"
repeat_region
complement(1168..1455)
/rpt_family="Alu-Sg"
repeat_region
1870..2157
/rpt_family="Alu-Sxz"
repeat_region
complement(3106..3396)
/rpt_family="Alu-Sx"
repeat_region
3563..3735
/rpt_family="MIR"
repeat_region
complement(4063..4355)
/rpt_family="Alu-Sx"
repeat_region
4694..4925
/rpt_family="L1MB3"
repeat_region
complement(4933..5218)
/rpt_family="Alu-Jo"
repeat_region
5380..5677
/rpt_family="Alu-Sz"
repeat_region
7958..8247
/rpt_family="Alu-Sz"
repeat_region
8531..8820
/rpt_family="Alu-Jo"
repeat_region
complement(8930..9219)
/rpt_family="Alu-Sx"
repeat_region
9542..9817
/rpt_family="Alu-Sg"
repeat_region
9861..10154
/rpt_family="Alu-Jo"
repeat_region
11127..11388
/rpt_family="Alu-Sx"
repeat_region
12278..12567
/rpt_family="Alu-Jo"
repeat_region
complement(12856..13166)
/rpt_family="Alu-Jo"
repeat_region
13866..13994
/rpt_family="Alu-Jo"
repeat_region
14003..14289
/rpt_family="Alu-Sx"
repeat_region
14505..14791
/rpt_family="Alu-Y"
repeat_region
complement(15165..15307)
/rpt_family="Alu-Jo"
repeat_region
15829..15960
/rpt_family="Alu-Spgxz"
repeat_region
15961..16259
/rpt_family="Alu-Sz"
repeat_region
16262..16333
/rpt_family="L1"
repeat_region
16336..16451
/rpt_family="L1MB3"
repeat_region
16452..16724

```



```

/mol type="genomic DNA"
/db xref="taxon:9606"
/chrnosome="16"
/clone="RP11-486111"

ORIGIN
Query Match          76.8%; Score 19.2; DB 8; Length 148412;
Best Local Similarity 87.5%; Pred. No. 1.5e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAACTCTGTAGGCTTTCAGT 24
    ||||| ||||| ||||| ||||| |||||
Db 146701 AAAAAATTCGTAGGTTTTCATT 146724

RESULT 42
LOCUS CS086332 150525 bp DNA linear PAT 25-MAY-2005
DEFINITION Sequence 36 from Patent WO2005042786.
ACCESSION CS086332
VERSION CS086332.1 GI:66711904
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.

REFERENCE
1 Harris, C. and Davis, L.
AUTHORS Compositions and methods for glioma classification
TITLE Patent: WO 2005042786-A 36 12-MAY-2005;
JOURNAL Exagen Diagnostics, Inc. (US)
FEATURES
    source
        1. 150525
            /organism="Homo sapiens"
            /mol_type="unassigned DNA"
            /db_xref="taxon:9606"

ORIGIN
Query Match          76.8%; Score 19.2; DB 6; Length 150525;
Best Local Similarity 87.5%; Pred. No. 1.5e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAACTCTGTAGGCTTTCAGT 24
    ||||| ||||| ||||| ||||| |||||
Db 92490 AAAAAATTCGTAGGTTTTCATT 92513

RESULT 43
AC107677 157746 bp DNA linear ROD 05-OCT-2004
LOCUS Mus musculus chromosome 3, clone RP23-235F22, complete sequence.
DEFINITION AC107677
ACCESSION AC107677.7 GI:53793869
VERSION HTG.
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidae; Muridae; Murinae; Mus.

REFERENCE
1 (bases 1 to 157746)
AUTHORS Birren, B., Nusbaum, C. and Lander, E.
TITLE Mus musculus chromosome 3, clone RP23-235F22
JOURNAL Unpublished
REFERENCE
2 (bases 1 to 157746)
AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,
Anderson, S., Barna, N., Bastien, V., Boguslavskiy, L., Boukhgalter, B.,
Brown, A., Camarata, J., Campolano, A., Chang, J., Chazaro, B.,
Coepe, P., Colangelo, M., Collins, S., Collymore, A., Cook, A.,
Cooke, P., DeArellano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S.,
Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S.,
Ginde, S., Gord, S., Goyette, M., Graham, L., Horton, L., Hulme, W.,
Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,

```

```

Kamat, A., Karatas, A., Kells, C., LaRocque, K., Lamazares, R.,
Landers, T., Lehoczy, J., Levine, R., Liu, G., Maclean, C.,
Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M.,
McEwan, P., McKernan, K., Meldrim, J., Meneus, L., Mihova, T.,
Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C.,
Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J.,
Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Roman, C.,
Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Royman, J.,
Rosetti, M., Roy, A., Santos, R., Schauer, S., Schuback, R., Seaman, S.,
Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Strauss, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J.,
Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (24-JAN-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 157746)
Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N.,
Anderson, M., Anderson, S., Atachchi, H.M., Barna, N., Bastien, V.,
Bloom, T., Boguslavskiy, L., Boukhgalter, B., Camarata, J., Chang, J.,
Coepe, P., Collymore, A., Cook, A., Cooke, P., Corum, B.,
DeArellano, K., Diaz, J. S., Dodge, S., Dooley, K., Dorris, L.,
Erickson, J., Faro, S., Ferreira, P., FitzGerald, M., Gage, D.,
Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N.,
Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I.,
Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T.,
Levine, R., Lindblad-Toh, K., Liu, G., Liu, X., Lui, A., Mabbitt, R.,
Maclean, C., Macdonald, P., Major, J., Manning, J., Matthews, C.,
McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V.,
Murphy, T., Naylor, J., Nguyen, C., Nguyen, T., Nicol, R., Norbu, C.,
O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K.,
Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C.,
Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schuback, R.,
Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N.,
Stojanovic, N., Stubbs, M., Talamas, J., Testaye, S., Theodore, J.,
Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R.,
Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L.,
Zimmer, A. and Zody, M.

Direct Submission
Submitted (11-SEP-2004) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 157746)
Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N.,
Anderson, M., Anderson, S., Atachchi, H.M., Barna, N., Bastien, V.,
Bloom, T., Boguslavskiy, L., Boukhgalter, B., Camarata, J., Chang, J.,
Coepe, P., Collymore, A., Cook, A., Cooke, P., Corum, B.,
DeArellano, K., Diaz, J. S., Dodge, S., Dooley, K., Dorris, L.,
Erickson, J., Faro, S., Ferreira, P., FitzGerald, M., Gage, D.,
Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N.,
Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I.,
Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T.,
Levine, R., Lindblad-Toh, K., Liu, G., Liu, X., Lui, A., Mabbitt, R.,
Maclean, C., Macdonald, P., Major, J., Manning, J., Matthews, C.,
McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V.,
Murphy, T., Naylor, J., Nguyen, C., Nguyen, T., Nicol, R., Norbu, C.,
O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K.,
Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C.,
Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schuback, R.,
Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N.,
Stojanovic, N., Stubbs, M., Talamas, J., Testaye, S., Theodore, J.,
Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R.,
Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L.,
Zimmer, A. and Zody, M.

Direct Submission
Submitted (05-OCT-2004) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Oct 5, 2004 this sequence version replaced gi:51988101.
All repeats were identified using RepeatMasker:
Smith, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/MIT Center for Genome Research

```

Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@broad.mit.edu
----- Project Information
Center project name: LI7624
Center clone name: 235_F_22

```
FEATURES
Source
1..157746
    Location/Qualifiers
        clone_lib="RP23-235F22"
        clone_lib="RP23-235F22"
        complement(243..343)
        /rpt_family="PB1D7"
        1830..1897
        /rpt_family="(TA)n"
        complement(2283..2453)
        /rpt_family="B2_Mm1"
        complement(2509..2772)
        /rpt_family="L1_MM"
        2772..3096
        /rpt_family="Lx"
        complement(3734..3904)
        /rpt_family="URR1A"
        3921..4115
        /rpt_family="B4A"
        4154..4223
        /rpt_family="(CATA)n"
        5473..5640
        /rpt_family="B3A"
        6264..6445
        /rpt_family="(TATG)n"
        complement(6449..6718)
        /rpt_family="B4"
        7602..7737
        /rpt_family="B1_MM"
        7738..7776
        /rpt_family="(CAAAA)n"
        8885..8935
        /rpt_family="AT_Rich"
        complement(8936..9161)
        /rpt_family="Lx2"
        9164..10365
        /rpt_family="Lx2"
        10367..10403
        /rpt_family="L1_Cp"
        11094..11172
        /rpt_family="(CAAAA)n"
        11872..11909
        /rpt_family="(TG)n"
        complement(11913..12056)
        /rpt_family="RSINE1"
        complement(12090..12475)
        /rpt_family="LIM2"
        12491..12528
        /rpt_family="(TCC)n"
        complement(12720..13181)
        /rpt_family="LIM2"
        13609..13760
        /rpt_family="(CACCC)n"
        15181..15513
        /rpt_family="RLTRETN_MM"
        15986..16050
        /rpt_family="AT_Rich"
        16719..16873
        /rpt_family="Zaphod"
        18776..18806
        /rpt_family="(TAAA)n"
        19190..19253
```

```
repeat_region      /rpt_family="(CA)n"
19718..20155
/rpt_family="Lx"
21103..21138
/rpt_family="(TTTG)n"
21471..21582
/rpt_family="MIR3"
complement(21875..22073)
/rpt_family="B3"
22128..22152
/rpt_family="AT_Rich"
23624..23668
/rpt_family="(TAAA)n"
24248..24346
/rpt_family="Lx9"
complement(25111..25342)
/rpt_family="RLTR11A"
25518..25794
/rpt_family="L1_MM"
26204..28475
/rpt_family="Lx4"
28496..28573
/rpt_family="(GA)n"
28574..29813
/rpt_family="Lx4"
complement(30658..30949)
/rpt_family="B4"
complement(31119..31258)
/rpt_family="B4"
complement(31406..31730)
/rpt_family="MLT1C"

Query Match      76.8%; Score 19.2; DB 9; Length 157746;
Best Local Similarity 87.5%; Pred.No.1.5e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 AAAAACTCTGTAGGCTTTTCAGTG 25
|||||
Db 111564 AAAGACTCTGTAGGCTGTAGTG 111587

RESULT 44
AC007226/c
LOCUS      159613 bp DNA linear PRI 02-DRC-2000
DEFINITION Homo sapiens chromosome 16 clone RPCI-11_520B23, complete sequence.
ACCESSION AC007226
VERSION AC007226.3 GI:11527452
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Homiidae; Homo.
REFERENCE 1 (bases 1 to 159613)
AUTHORS Bruce,D., Mundt,M., Doggett,N., Munk,C., Saunders,E., Robinson,D.,
            Jones,M., Buckingham,J., Chasteen,L., Thompson,S., Goodwin,L.,
            Bryant,J., Tesmer,J., Meincke,L., Longmire,J., White,S., Tatum,O.,
            Campbell,C., Fawcett,J., Maltbie,M., Bussod,M., Sutherland,R.,
            Han,C. and Deaven,L.
TITLE      Sequencing of Human Chromosome 16q12
JOURNAL    Unpublished
REFERENCE 2 (bases 1 to 159613)
AUTHORS Ricke,D.O.
TITLE      Large Scale Sequence Analysis and Annotation with the Sequence
            Comparison Analysis (SCAN) System
JOURNAL    Unpublished
REFERENCE 3 (bases 1 to 159613)
AUTHORS Bruce,D., Mundt,M., Doggett,N., Munk,C., Saunders,E., Robinson,D.,
            Jones,M., Buckingham,J., Chasteen,L., Thompson,S., Goodwin,L.,
            Bryant,J., Tesmer,J., Meincke,L., Longmire,J., White,S., Tatum,O.,
            Campbell,C., Fawcett,J., Maltbie,M., Bussod,M., Sutherland,R.,
            McMurry,K., Han,C. and Deaven,L.
TITLE      Direct Submission
```



```
JOURNAL      Submitted (06-APR-1999) Center for Human Genome Studies, DOE Joint
              Genome Institute, Los Alamos National Laboratory, MS M888, Los
REFERENCE     Alamos, NM 87545, USA
AUTHORS       4 (bases 1 to 159613)
              Bruce, D., Mundt, M., Doggett, N., Munk, C., Saunders, E., Robinson, D.,
              Jones, M., Buckingham, J., Chasteen, L., Thompson, S., Goodwin, L.,
              Bryant, J., Temer, J., Meincke, L., Longmire, J., White, S., Tatum, O.,
              Campbell, C., Fawcett, J., Maltbie, M., Bussod, M., Sutherland, R.,
              McMurtry, K., Han, C. and Deaven, L.
TITLE         Direct Submission
JOURNAL       Submitted (19-JAN-2000) Center for Human Genome Studies, DOE Joint
              Genome Institute, Los Alamos National Laboratory, MS M888, Los
REFERENCE     Alamos, NM 87545, USA
AUTHORS       5 (bases 1 to 159613)
              Mundt, M.O.
TITLE         Direct Submission
JOURNAL       Submitted (02-DEC-2000)
COMMENT       On Dec 2, 2000 this sequence version replaced gi:6715702.
              overlaps RPCI-11.49006.
FEATURES     Location/Qualifiers
              source          1..159613
              /organism="Homo sapiens"
              /mol_type="genomic DNA"
              /db_xref="taxon:9606"
              /chromosomes="16"
              /clone="RPCI-11_520B23"
              628..1000
              /rpt_family="Alu"
              repeat_region   914..1106
              repeat_region   /rpt_family="Alu"
              misc_feature    1381..1433
              /note="GRAIL 2 excellent exon, frame 2"
              repeat_region   complement(2941..3231)
              repeat_region   /rpt_family="Alu"
              repeat_region   6999..7300
              repeat_region   /rpt_family="Alu"
              repeat_region   7219..7619
              repeat_region   /rpt_family="Alu"
              repeat_region   7994..8323
              repeat_region   /rpt_family="Alu"
              repeat_region   8754..8812
              repeat_region   /rpt_family="L1"
              repeat_region   complement(9682..9958)
              repeat_region   /rpt_family="Alu"
              repeat_region   10288..10585
              repeat_region   /rpt_family="Alu"
              repeat_region   10825..11133
              repeat_region   /rpt_family="Alu"
              repeat_region   /rpt_family="Alu"
              repeat_region   complement(11595..11675)
              repeat_region   /rpt_family="MIR"
              repeat_region   complement(12225..12388)
              repeat_region   /rpt_family="Alu"
              repeat_region   complement(12521..12831)
              repeat_region   /rpt_family="Alu"
              repeat_region   complement(12847..13106)
              repeat_region   /rpt_family="Alu"
              repeat_region   13339..14071
              repeat_region   /rpt_family="MER7"
              repeat_region   complement(13414..13746)
              repeat_region   /rpt_family="GOLEM"
              repeat_region   complement(14064..14705)
              repeat_region   /rpt_family="Alu"
              repeat_region   14716..15019
              repeat_region   /rpt_family="Alu"
              repeat_region   14824..15149
              repeat_region   /rpt_family="Alu"
              repeat_region   15254..16191
              repeat_region   /rpt_family="MER7"
              repeat_region   complement(15376..15455)
              repeat_region   /rpt_family="GOLEM"
              repeat_region   15492..15972
              repeat_region   /rpt_family="Alu"
              repeat_region   complement(16126..16182)
              repeat_region   /rpt_family="Alu"
              repeat_region   complement(16974..17125)
              repeat_region   /rpt_family="Alu"
              repeat_region   17363..17660
              repeat_region   /rpt_family="Alu"
              repeat_region   17810..18125
              repeat_region   /rpt_family="Alu"
              repeat_region   18224..18546
              repeat_region   /rpt_family="Alu"
              repeat_region   complement(18702..18838)
              repeat_region   /rpt_family="MLT1"
              repeat_region   complement(18786..18935)
              repeat_region   /rpt_family="MLT1"
              repeat_region   complement(19030..19328)
              repeat_region   /rpt_family="Alu"
              repeat_region   20312..20480
              repeat_region   /rpt_family="Alu"
              repeat_region   20510..21865
              repeat_region   /rpt_family="Alu"
              repeat_region   20863..21159
              repeat_region   /rpt_family="Alu"
              repeat_region   21045..21405
              repeat_region   /rpt_family="Alu"
              repeat_region   21356..21700
              repeat_region   /rpt_family="Alu"
              repeat_region   complement(21877..22074)
              repeat_region   /rpt_family="MER30"
              repeat_region   complement(22009..22397)
              repeat_region   /rpt_family="Alu"
              misc_feature    join(22997..23104,24088..24231,26704..26889,27666..27815,
              27905..28030,32507..32602,32686..32823,34524..34817,
              36396..36701,36460..36801)
              /note="CAA07491.1"
              /function="PBK1 protein (Homo sapiens)"
              23520..23718
              /rpt_family="Alu"
              repeat_region   24383..24662
              repeat_region   /rpt_family="Alu"
              repeat_region   complement(25105..25404)
              repeat_region   /rpt_family="Alu"
              repeat_region   25430..25996
              repeat_region   /rpt_family="Alu"
              repeat_region   complement(26232..26539)
              repeat_region   /rpt_family="Alu"
              repeat_region   complement(27125..27401)
              repeat_region   /rpt_family="Alu"
              repeat_region   complement(28086..28402)
              repeat_region   /rpt_family="Alu"
              repeat_region   28477..28786
              repeat_region   /rpt_family="Alu"
              repeat_region   28803..28988
              repeat_region   /rpt_family="MER3"
              repeat_region   complement(29093..29338)
              repeat_region   /rpt_family="Alu"
              repeat_region   complement(30620..30923)
              repeat_region   /rpt_family="Alu"
              repeat_region   31139..31747
              repeat_region   /rpt_family="Alu"
              repeat_region   31836..32118
              repeat_region   /rpt_family="Alu"
              repeat_region   complement(32156..32378)
              repeat_region   /rpt_family="Alu"
              repeat_region   complement(33092..33381)
              repeat_region   /rpt_family="Alu"
              repeat_region   complement(33409..33562)
              repeat_region   /rpt_family="MER6"
              repeat_region   complement(33932..34239)
              repeat_region   /rpt_family="Alu"
              repeat_region   complement(34945..35482)
              repeat_region   /rpt_family="Alu"
              repeat_region   complement(35626..35930)
              repeat_region   /rpt_family="Alu"
              repeat_region   36030..36329
```


JOURNAL Submitted (22-JUN-2005) Broad Institute of MIT and Harvard, 320 Charles Street, Cambridge, MA 02141, USA
 COMMENT On Jun 22, 2005 this sequence version replaced gi:67763997. All repeats were identified using RepeatMasker: Smit, A.F.A. & Green, P. (1996-1997) <http://ftp.genome.washington.edu/RM/RepeatMasker.html>
 ----- Genome Center
 Center: Broad Institute of MIT and Harvard
 Center code: WBR
 Web site: <http://www-seq.wi.mit.edu>
 Contact: sequence_submissions@broad.mit.edu
 ----- Project Information
 Center project name: L32316
 Center clone name: 273_M_3

* NOTE: This is a 'working draft' sequence. It currently consists of 4 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

* 1 21232: contig of 21232 bp in length
 * 21233: gap of unknown length
 * 21333: contig of 69983 bp in length
 * 91316: gap of unknown length
 * 91415: gap of unknown length
 * 145954: contig of 54539 bp in length
 * 146054: gap of unknown length
 * 146055: 161009: contig of 14955 bp in length.

FEATURES
 source
 1. .161009 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /chromosome="8"
 /map="8"
 /clone="RP24-273M3"
 /clone_lib="RPCI-24 Male Mouse BAC"
 21233. .21332 /estimated_length=unknown
 91316. .91415 /estimated_length=unknown
 145955. .146054 /estimated_length=unknown

ORIGIN
 Query Match 76.8% Score 19.2; DB 14; Length 161009;
 Best Local Similarity 87.5%; Pred. No. 1.5e+02;
 Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAATCTCTGTAGGCTTTCAGTG 25
 ||||| ||| ||||| ||||| |||||
 Db 69604 AAAACAACTGAGGCTTTCAGTG 69627

RESULT 47
 AP000356/c 163795 bp DNA linear PRI 16-JUN-1999
 LOCUS Homo sapiens genomic DNA, chromosome 22q11.2, clone KB1995A5.
 DEFINITION AP000356
 ACCESSION AP000356
 VERSION AP000356.1 GI:5103019
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 163795)
 Shimizu,N.
 TITLE Human DNA sequence from clone KB1995A5 on chromosome 22q11.2
 JOURNAL Published Only in Database (1999)
 REFERENCE 2 (bases 1 to 163795)

AUTHORS Shimizu,N.
 TITLE Direct Submission
 JOURNAL Submitted (10-JUN-1999) Nobuyoshi Shimizu, Keio university, school of medicine, Molecular Biology; 35 Shinanomachi, Shinjuku-ku, Tokyo 160-0016, Japan (E-mail:shimizu@mb.med.keio.ac.jp, Tel:81-3-3351-2370, Fax:81-3-3351-2370)
 COMMENT This is a complete sequence of the insert of KB1995A5 clone. The proximal adjacent clone is KB1896H10 (Acc.#AP000355) with 28999-bp overlapping. The distal adjacent clone is N95F10 (Acc.#AP000357) with 9742-bp overlapping.
 FEATURES
 Location/Qualifiers
 1. .163795
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="22"
 /map="22q11.2"
 /clone="KB1995A5"
 601. .1212 /rpt_family="ALU"
 2262. .2556 /rpt_family="ALU"
 complement(2631. .2808)
 /rpt_family="L1"
 complement(3926. .4211)
 /rpt_family="ALU"
 complement(5143. .5442)
 /rpt_family="L1"
 8009. .8291 /rpt_family="ALU"
 complement(9028. .9324)
 /rpt_family="ALU"
 10207. .10469 /rpt_family="ALU"
 13212. .13500 /rpt_family="ALU"
 13522. .13782 /rpt_family="ALU"
 14972. .15263 /rpt_family="ALU"
 complement(16821. .17054)
 /rpt_family="ALU"
 19386. .19669 /rpt_family="ALU"
 20189. .20429 /rpt_family="ALU"
 21005. .21126 /rpt_family="ALU"
 /rpt_family="ALU"
 21278. .21587 /rpt_family="ALU"
 complement(21639. .21859)
 /rpt_family="ALU"
 22062. .22294 /rpt_family="ALU"
 22353. .22493 /rpt_family="L1"
 22667. .22920 /rpt_family="ALU"
 complement(23138. .23416)
 /rpt_family="ALU"
 complement(24062. .24373)
 /rpt_family="ALU"
 24825. .25114 /rpt_family="ALU"
 complement(25386. .25553)
 /rpt_family="ALU"
 complement(25849. .25975)
 /rpt_family="ALU"
 26752. .27137 /rpt_family="ALU"
 complement(27233. .27532)
 /rpt_family="ALU"
 27970. .28250 /rpt_family="ALU"

```
repeat_region /rpt_family="ALU"
29418. .29707
repeat_region /rpt_family="ALU"
complement(30743. .31449)
repeat_region /rpt_family="ALU"
complement(31594. .31881)
repeat_region /rpt_family="ALU"
32106. .32347
/rpt_family="ALU"
33680. .33945
repeat_region /rpt_family="ALU"
37041. .37337
repeat_region /rpt_family="ALU"
complement(37911. .38207)
repeat_region /rpt_family="ALU"
complement(38286. .38560)
repeat_region /rpt_family="ALU"
complement(41817. .42107)
repeat_region /rpt_family="ALU"
44801. .45082
/rpt_family="ALU"
complement(45132. .45223)
repeat_region /rpt_family="ALU"
45524. .45614
/rpt_family="MER30"
46248. .46442
repeat_region /rpt_family="MER"
complement(48671. .48806)
repeat_region /rpt_family="ALU"
49365. .49531
/rpt_family="MER3"
complement(52097. .52183)
repeat_region /rpt_family="MLT1"
complement(52391. .52711)
repeat_region /rpt_family="ALU"
complement(52745. .53004)
repeat_region /rpt_family="ALU"
complement(53028. .53327)
repeat_region /rpt_family="ALU"
complement(53372. .53689)
repeat_region /rpt_family="ALU"
54030. .54282
/rpt_family="ALU"
complement(54365. .54437)
repeat_region /rpt_family="ALU"
54541. .55027
/rpt_family="ALU"
55398. .55449
repeat_region /rpt_family="ALU"
complement(56323. .56399)
repeat_region /rpt_family="MER42"
complement(57185. .57462)
repeat_region /rpt_family="ALU"
complement(57966. .58242)
repeat_region /rpt_family="ALU"
complement(58582. .58851)
repeat_region /rpt_family="ALU"
59776. .60056
/rpt_family="ALU"
complement(60648. .60821)
repeat_region /rpt_family="ALU"
complement(60861. .61300)
repeat_region /rpt_family="ALU"
61779. .62664
/rpt_family="MER21"
62354. .62627
/rpt_family="MER39"
62703. .62991
repeat_region /rpt_family="ALU"
complement(63120. .63407)
repeat_region /rpt_family="ALU"
63780. .64061
/rpt_family="ALU"
```

```
repeat_region 64794. .65072
/rpt_family="ALU"
65255. .65482
/rpt_family="ALU"
65910. .66157
/rpt_family="ALU"
complement(66844. .67133)
/rpt_family="ALU"
67645. .67943
/rpt_family="ALU"
complement(68043. .68339)
/rpt_family="ALU"
complement(68550. .68851)
/rpt_family="ALU"
69601. .69888
/rpt_family="ALU"
69979. .70259
/rpt_family="ALU"
71892. .72253
/rpt_family="L1"
72368. .72636
/rpt_family="ALU"
73230. .73512
/rpt_family="ALU"
73847. .74133
/rpt_family="ALU"
complement(75026. .75305)
/rpt_family="ALU"
complement(78843. .79128)
/rpt_family="ALU"
complement(82429. .82707)
/rpt_family="ALU"
complement(83392. .83666)
/rpt_family="ALU"
85767. .86104
/rpt_family="ALU"
complement(88527. .88684)
/rpt_family="MLT1"
complement(88766. .88888)
repeat_region
```

Query Match 76.8%; Score 19.2; DB 8; Length 163795;

Best Local Similarity 87.5%; Pred. NO. 1.5e+02;

Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAATCTCTAGGCTTTCAGTG 25

Db 70436 AAAAATCTCTAGCTTTGATTG 70413

RESULT 48

AC009048/c

LOCUS

DEFINITION

AC009048

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

AC009048 163844 bp DNA linear PRI 28-JUN-2001
Homo sapiens chromosome 16 clone RP11-211H13, complete sequence.

AC009048

AC009048

AC009048.6 GI:14573673

HTG. Homo sapiens (human)

Source

Organism

Reference

Authors

Title

Journal

Reference

Authors

Title

Journal

Reference

Authors

Title

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

1 (bases 1 to 163844)

DOE Joint Genome Institute and Stanford Human Genome Center.

Direct Submission

Unpublished

2 (bases 1 to 163844)

DOE Joint Genome Institute.

Direct Submission

Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA

3 (bases 1 to 163844)

DOE Joint Genome Institute and Stanford Human Genome Center.

Direct Submission

JOURNAL Submitted (28-JUN-2001) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA

COMMENT On Jun 28, 2001 this sequence version replaced gi:7689878. Draft Sequence Produced by DOE Joint Genome Institute www.jgi.doe.gov

Finishing Completed at Stanford Human Genome Center www.shgc.stanford.edu

Quality: Phrap Quality >40 99.7% of Sequence; Estimated Total Number of Errors is 0.4.

STS Content: SHGC-52683 G36700 SHGC-60499 G36838.

FEATURES
source Location/Qualifiers
1..163844
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="16"
/clone="RP11-211H13"

ORIGIN
Query Match 76.8%; Score 19.2; DB 8; Length 163844;
Best Local Similarity 87.5%; Pred. No. 1.5e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAACTCCTGTAGGCTTTTCAGT 24
|||||
Db 59236 AAAAAATCCTGTAGGCTTTTCATT 59213
|||||

RESULT 49
AC019081/c
LOCUS AC019081.1 GI:6648469
DEFINITION HTG; HTGS_PHASE1.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 163865)
Waterston,R.H.
TITLE The sequence of Homo sapiens clone
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 163865)
Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (30-DEC-1999) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA

COMMENT Center project name: H NH0340P24.
* NOTE: This is a 'working draft' sequence. It currently consists of 11 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 2004: contig of 2004 bp in length
2005 2022: gap of unknown length
2023 5849: contig of 3827 bp in length
5850 5867: gap of unknown length
5868 9538: contig of 3671 bp in length
9539 9556: gap of unknown length
16258 16276: gap of unknown length
16277 22543: contig of 6267 bp in length
22544 22561: gap of unknown length
22562 29557: contig of 6996 bp in length

JOURNAL Submitted (28-JUN-2001) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA

COMMENT On Jun 28, 2001 this sequence version replaced gi:7689878. Draft Sequence Produced by DOE Joint Genome Institute www.jgi.doe.gov

Finishing Completed at Stanford Human Genome Center www.shgc.stanford.edu

Quality: Phrap Quality >40 99.7% of Sequence; Estimated Total Number of Errors is 0.4.

STS Content: SHGC-52683 G36700 SHGC-60499 G36838.

FEATURES
source Location/Qualifiers
1..163865
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="RP11-340P24"

ORIGIN
Query Match 76.8%; Score 19.2; DB 14; Length 163865;
Best Local Similarity 87.5%; Pred. No. 1.5e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAACTCCTGTAGGCTTTTCAGT 25
|||||
Db 147097 AAAAACTCCTGTAGGCTTTTGAGG 147074
|||||

RESULT 50
AC147169/c
LOCUS AC147169
DEFINITION Xenopus tropicalis clone CH216-149115, complete sequence.
ACCESSION AC147169
VERSION AC147169.2 GI:54606892
KEYWORDS HTG.
SOURCE Xenopus tropicalis (Silurana tropicalis)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Xenopodinae; Xenopus; Silurana.
REFERENCE 1 (bases 1 to 165948)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 165948)
DOE Joint Genome Institute.
AUTHORS Direct Submission
TITLE Direct Submission
JOURNAL Submitted (05-NOV-2003) Production Genomics Facility, DOE Joint Genome Institute, 2800 Mitchell Drive B100, Walnut Creek, CA 94598-1698, USA
REFERENCE 3 (bases 1 to 165948)
Stanford Human Genome Center.
AUTHORS DOE Joint Genome Institute
CONSTRM Direct Submission
TITLE Direct Submission

JOURNAL Submitted (23-OCT-2004) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT On Oct 23, 2004 this sequence version replaced gi:38175822.
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www.sbgc.stanford.edu

Quality: Phrap Quality >=40 100% of Sequence;
Estimated Total Number of Errors is 0.

FEATURES
source
1..165948
Location/Qualifiers
/organism="Xenopus tropicalis"
/mol_type="genomic DNA"
/db_xref="taxon:8364"
/clone="CH216-149115"

ORIGIN

Query Match 76.8%; Score 19.2; DB 5; Length 165948;
Best Local Similarity 87.5%; Pred No. 1.5e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 AAAAACTCCTGTAGGCTTTTCAGT 24
||||| ||||| ||||| |||||
Db 36239 AAAAACTCCTGTAGGCTGCCAGT 36216

Search completed: February 3, 2006, 21:26:53
Job time : 875 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 3, 2006, 10:53:33 ; Search time 203.111 Seconds
(without alignments)
820.326 Million cell updates/sec

Title: US-10-719-900-20

Perfect score: 25

Sequence: 1 aaaaactcctgtaggtttcagt 25

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database : N Geneseq 21.*

- 1: Geneseqn1980s.*
- 2: Geneseqn1990s.*
- 3: Geneseqn2000s.*
- 4: Geneseqn2001as.*
- 5: Geneseqn2001bs.*
- 6: Geneseqn2002as.*
- 7: Geneseqn2002bs.*
- 8: Geneseqn2003as.*
- 9: Geneseqn2003bs.*
- 10: Geneseqn2003cs.*
- 11: Geneseqn2003ds.*
- 12: Geneseqn2004as.*
- 13: Geneseqn2004bs.*
- 14: Geneseqn2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	25	100.0	74868	13	ADV34991 Murine CD
C 2	19.8	79.2	599	13	Adq55481 Novel can
C 3	19.8	79.2	2243	13	Adt99098 GI to S p
C 4	19.2	76.8	439	10	Abt22339 Breast ca
C 5	19.2	76.8	545	13	Adul4171 Solid tum
C 6	19.2	76.8	562	4	Aak12378 Human bra
C 7	19.2	76.8	562	4	Abt37712 Human liv
C 8	19.2	76.8	562	6	Abt37712 Human liv
C 9	19.2	76.8	2527	11	Adm02046 Human gen
C 10	19.2	76.8	2587	10	Adc35122 Human bre
C 11	19.2	76.8	2587	12	Adn03792 Antipsori
C 12	19.2	76.8	2587	12	Adp13389 Renal cel
C 13	19.2	76.8	2587	13	Adp54303 Human PRO
C 14	19.2	76.8	2587	13	Adp24764 PRO poly
C 15	19.2	76.8	2587	13	Adu06152 Novel bro
C 16	19.2	76.8	2587	14	Ady20501 DNA encod
C 17	19.2	76.8	2587	14	Ady14680 DNA encod
C 18	19.2	76.8	3866	10	Ades3710 Human pro
C 19	19.2	76.8	5569	5	Adl61917 Human ova

Abv24840	Human pro	5570	76.8	19.2	5	ABV24840
Aak84729	Human imm	48037	76.8	19.2	4	Aak84729
Aak85983	Human imm	48037	76.8	19.2	4	Aak85983
Aak84730	Human imm	48045	76.8	19.2	4	Aak84730
Aak85984	Human imm	48045	76.8	19.2	4	Aak85984
Ad433112	Human mit	466	74.4	18.6	10	AD433112
Ad48290	Human enz	39982	74.4	18.6	8	AD48290
Adl13771	Osteoearth	201766	74.4	18.6	10	ADL13771
Aac36832	Arabidops	1175	73.6	18.4	3	AAC36832
Aac42714	Arabidops	2286	73.6	18.4	3	AAC42714
Adn74670	Thale cre	2286	73.6	18.4	3	ADN74670
Abn22936	Human ORF	463	72.8	18.2	6	ABN22936
Continuation (2 of		110000	72.8	18.2	10	ADH10017_1
Adq97266	Human can	110000	72.8	18.2	12	ADQ97266_0
Continuation (2 of		110000	72.8	18.2	12	ADQ97266_1
Continuation (3 of		110000	72.8	18.2	12	ADQ97328_2
Abd33491	Human can	112241	72.8	18.2	13	ABD33491_2
Adl13676	Osteoearth	192839	72.8	18.2	10	ADL13676
Aan82201	Beta-amyl	1794	71.2	17.8	1	AAN82201
Continuation (4 of		110000	71.2	17.8	6	ABQ69245_03
Continuation (3 of		110000	71.2	17.8	6	ABQ67197_02
Aal47655	Specific	43	70.4	17.6	6	AAL47655_03
Aal47656	Specific	43	70.4	17.6	6	AAL47656_03
Adb50438	Primary r	457	70.4	17.6	10	ADB50438
Adc56904	Bacterial	925	70.4	17.6	13	ADC56904
Adx09691	Plant ful	1303	70.4	17.6	13	ADX09691
Ab11811	Drosophil	1368	70.4	17.6	4	AB11811
Aag99349	Frog bomb	1563	70.4	17.6	2	AAG99349
Aas12704	Frog bomb	1563	70.4	17.6	5	AAS12704
Ades83485	Human gen	2962	70.4	17.6	10	ADE83485
Ab11810	Drosophil	5904	70.4	17.6	4	AB11810
Ab118535	Drosophil	8685	70.4	17.6	4	AB118535
Adz13273	Human can	99544	70.4	17.6	14	ADZ13273
Acn45034	Human gen	99588	70.4	17.6	11	ACN45034
Ado34435	Human SLI	110000	70.4	17.6	12	ADO34435_0
Continuation (2 of		110000	70.4	17.6	12	ADO34435_1
Abk83497	Human CDN	227968	70.4	17.6	6	ABK83497
Adq18538	Human sof	227968	70.4	17.6	12	ADQ18538
Adg97146	Human can	348101	70.4	17.6	12	ADG97146
Aak54030	Murine tr	356	68.8	17.2	4	AAK54030
Achl6777	Human adu	426	68.8	17.2	9	ACH16777
Adg79472	Novel can	538	68.8	17.2	13	ADG79472
Abn60093	Human can	576	68.8	17.2	6	ABN60093
Adc163078	Human col	627	68.8	17.2	14	ADC163078
Aah03914	Human CDN	781	68.8	17.2	4	AAH03914
Aah13859	Human CDN	1233	68.8	17.2	4	AAH13859
Aca57585	Human adi	1293	68.8	17.2	8	ACA57585
Aah33825	Human col	1315	68.8	17.2	4	AAH33825
Adq20266	Human sof	1348	68.8	17.2	12	ADQ20266
Ad35833	Novel mou	1425	68.8	17.2	4	AD35833
Aah98243	Human EST	1496	68.8	17.2	4	AAH98243
Adf76674	Prokaryot	1857	68.8	17.2	8	ADF76674
Adq86627	Human tum	1860	68.8	17.2	10	ADQ86627
Adq85502	Human tum	1884	68.8	17.2	13	ADQ85502
Adq24287	Human sof	1956	68.8	17.2	12	ADQ24287
Aai59067	Human pol	1971	68.8	17.2	4	AAI59067
Adq99290	DNA encod	1971	68.8	17.2	5	ADQ99290
Adb49050	Novel hum	1971	68.8	17.2	9	ADB49050
Adc138806	Rice stre	2000	68.8	17.2	11	ADC138806
Aas79442	DNA encod	2006	68.8	17.2	5	AAS79442
Aah17147	Human CDN	2008	68.8	17.2	4	AAH17147
Abn06762	Human CDN	2022	68.8	17.2	4	ABN06762
Aas31650	cDNA enco	2022	68.8	17.2	4	AAS31650
Abv84099	Human pol	2022	68.8	17.2	6	ABV84099
Adq64806	Novel hum	2386	68.8	17.2	12	ADQ64806
Aah16475	Human CDN	2491	68.8	17.2	4	AAH16475
Aak94708	Human ful	2491	68.8	17.2	4	AAK94708
Adl31708	Full leng	2794	68.8	17.2	12	ADL31708
Aai60853	Human pol	2794	68.8	17.2	4	AAI60853
Aac77585	Human ORF	2795	68.8	17.2	3	AAC77585
Aas60774	Human can	3376	68.8	17.2	4	AAS60774

c 93	17.2	68.8	3914	10	ADB62696	Adb62696 Human cDN
c 94	17.2	68.8	4263	3	AAa13379	Mouse Wt5
c 95	17.2	68.8	4263	3	AAa13379	Mouse Wt5
c 96	17.2	68.8	6138	3	AAa14341	CDNA enco
c 97	17.2	68.8	6138	3	AAa14341	CDNA enco
c 98	17.2	68.8	6141	6	ABV73471	Human cyt
c 99	17.2	68.8	6586	5	AAa45005	CDNA enco
c 100	17.2	68.8	7646	6	ABV73469	Human cyt
c 101	17.2	68.8	16570	6	ABV73470	Human cyt
c 102	17.2	68.8	108359	9	ADA13316	Human ffr
c 103	17.2	68.8	169998	6	ADA13316	Human ffr
c 104	17.2	68.8	197496	6	ABN85584	Human EGF
c 105	17.2	68.8	200620	12	ADO56277	Human pre
c 106	17.2	68.8	209083	13	ABD32854	Human can
c 107	17.2	68.8	213300	14	ADX80726	Human RAL
c 108	17.2	68.8	195	4	AAK83806	Human imm
c 109	17.2	68.8	199	4	AAK83806	Human imm
c 110	17.2	68.8	288	5	AAH52267	Human APP
c 111	17.2	68.8	412	11	ACL38127	Rice stre
c 112	17.2	68.8	419	5	AAa84713	DNA enco
c 113	17.2	68.8	425	5	AAa84710	DNA enco
c 114	17.2	68.8	430	9	ACH18048	Human adu
c 115	17.2	68.8	519	5	ABV14658	Human pro
c 116	17.2	68.8	549	8	ABZ19092	Group IIR
c 117	17.2	68.8	900	2	AAa67966	H. pylori
c 118	17.2	68.8	1090	13	ADR65002	Cotton cD
c 119	17.2	68.8	1095	10	ADH82740	Enterococ
c 120	17.2	68.8	1421	10	ADC32313	Human nov
c 121	17.2	68.8	1482	8	ACA48417	Prokaryot
c 122	17.2	68.8	1690	8	ABT42930	Human neu
c 123	17.2	68.8	1939	6	ABZ11684	Human pol
c 124	17.2	68.8	1939	12	ADM44202	Novel hum
c 125	17.2	68.8	1955	11	ADM03271	Human cDN
c 126	17.2	68.8	2000	12	ADJ41452	Plant cDN
c 127	17.2	68.8	2064	12	ADF88612	DNA of hu
c 128	17.2	68.8	2186	10	ADF88614	DNA of hu
c 129	17.2	68.8	2393	10	ADA52804	Human cod
c 130	17.2	68.8	2409	11	ADM01794	Human cDN
c 131	17.2	68.8	2526	14	ABE91776	DNA enco
c 132	17.2	68.8	2766	5	AAa91674	DNA enco
c 133	17.2	68.8	2776	10	ADE25649	Human cDN
c 134	17.2	68.8	2905	6	ABK84106	Human cDN
c 135	17.2	68.8	2905	8	ABZ23871	Human 2'-
c 136	17.2	68.8	2905	10	ADD18901	Human dis
c 137	17.2	68.8	2905	10	ADD18899	Human dis
c 138	17.2	68.8	2905	10	ADF76968	Novel hum
c 139	17.2	68.8	2905	12	ADJ75032	Marker ge
c 140	17.2	68.8	2905	12	ADN04093	Antipsori
c 141	17.2	68.8	2905	13	ADN25811	Breast ca
c 142	17.2	68.8	2905	13	ADR14236	Human NF-
c 143	17.2	68.8	2905	13	ADS64453	DNA enco
c 144	17.2	68.8	2905	14	ADY15451	DNA enco
c 145	17.2	68.8	2905	14	AEA23545	Human PRO
c 146	17.2	68.8	2905	14	AEA36131	Human nuc
c 147	17.2	68.8	2911	12	ADI32351	Human OAS
c 148	17.2	68.8	2911	12	ACN39568	Tumour-as
c 149	17.2	68.8	2921	8	ABZ35982	Human sec
c 150	17.2	68.8	3028	14	AEA61091	Human CLD

ALIGNMENTS

RESULT 1
ADV34991/c
ID ADV34991 standard; cDNA; 74868 BP.
XX
AC ADV34991;
XX
DT 10-FEB-2005 (first entry)
XX
DE Murine cDNA differentially expressed in the presence of valproate Seq67.
XX

RESULT 2
ADQ55481
ID ADQ55481 standard; DNA; 599 BP.
XX
AC ADQ55481;
XX
DT 21-OCT-2004 (first entry)
XX
DE Novel canine microarray-related DNA sequence SeqID6783.
XX

Query Match 100.0%; Score 25; DB 13; Length 74868;
Best Local Similarity 100.0%; Pred. No. 0.29; Indels 0; Gaps 0;
Matches 25; Conservative 0; Mismatches 0

QY 1 AAAAACTCCTGCTAGGCTTTCAGTG 25
Db 33482 AAAAACTCCTGCTAGGCTTTCAGTG 33458


```
XX canine microarray; drug screening; toxicity assay;
KW environmental pollutant; cellular response; gene expression profile;
KW toxic response; liver necrosis; fatty liver disease;
KW protein adduct formation; hepatitis; dog; ds.
XX
OS Canis familiaris.
XX WO2004063324-A2.
XX 29-JUL-2004.
XX
XX 05-MAY-2003; 2003WO-US013853.
XX
XX 03-MAY-2002; 2002US-0377240P.
XX
XX (GENE-) GENE LOGIC INC.
XX (PFIZ ) PFIZER PROD INC.
XX
XX Diggans JC, Porter M, Wei T;
XX WPI; 2004-561890/54.
XX
XX New isolated nucleic acid molecule, useful for drug screening and
PT toxicity assays or for assessing the impact, including toxicity, of a
PT compound, pharmaceutical agent or environmental pollutant on a cell or
PT living organism.
XX
PS Claim 1; SEQ ID NO 6783; 4lpp; English.
XX
CC This invention is related to a novel isolated canine nucleic acid
CC sequences and the construction of canine microarrays containing a
CC significant portion of the canine genome. The isolated canine nucleic
CC acid sequences of the invention may be useful for drug screening and
CC toxicity assays. The invention is therefore useful for assessing the
CC impact, including toxicity, of a compound, pharmaceutical agent or
CC environmental pollutant on a cell or living organism. The methods are
CC useful for detecting genes that are up- or down-regulated in canines in a
CC disease state. The sequences are useful as diagnostic agents or markers
CC to detect a cellular response in a sample individually or as part of a
CC gene expression profile. It is also useful as a target for agents that
CC modulate gene expression or activity. The database is useful for
CC producing electronic Northern blots that allow the user to determine the cell
CC type or tissue in which a given gene is expressed and to allow
CC determination of the abundance or expression level of a given gene in a
CC particular tissue or cell. The methods are useful for determining the
CC similarity of a toxic response to one or more individual compounds. The
CC methods are useful for predicting at least one toxic response or the
CC likelihood that a compound or test agent will induce various specific
CC pathologies such as those of the liver (liver necrosis, fatty liver
CC disease, protein adduct formation or hepatitis), those of the kidney,
CC heart, brain or testes, or other pathologies associated with at least one
CC of the toxins. The methods are also useful for predicting or elucidating
CC the potential cellular pathways influenced, induced or modulated by the
CC compound or test agent due to the similarity of the expression profile
CC compared to the profile induced by a known toxin. The present sequence is
CC that of a canine DNA sequence which was claimed for use during the
CC production of a canine microarray of the invention.
XX
SQ Sequence 599 BP; 182 A; 96 C; 99 G; 203 T; 0 U; 19 Other;
Query Match 79.2%; Score 19.8; DB 13; Length 599;
Best Local Similarity 91.3%; Pred. No. 39;
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 2 AAAAATCTCTGCTAGGCTTTTCAGT 24
Db 290 AAAAATCTCTGCTAGGCTATCAGT 312
RESULT 3
ADR99098/c
ID ADR99098 standard; DNA; 2243 BP.
```

```
XX ADR99098;
XX AC
XX 02-DEC-2004 (first entry)
XX DT
XX DE
XX G1 to S phase transition 1, GSPT1, coding sequence, SEQ ID 104.
XX KW Cytostatic; breast cancer; cancer; human; gene; ds; GSPT1.
XX OS Homo sapiens.
XX WO2004078035-A2.
XX 16-SEP-2004.
XX
XX 27-FEB-2004; 2004WO-US007268.
XX
XX 28-FEB-2003; 2003US-0450655P.
XX
XX (FARB ) BAYER PHARM CORP.
XX
XX Eveleigh D, Bigwood D;
XX WPI; 2004-553556/63.
XX P-PSDB; ADR99225.
XX GENBANK; BE906054.
XX
XX Diagnosing breast cancer comprises comparing the level of expression of
PT genes or gene products in a first biological sample taken from a patient
PT with that in a normal patient sample.
XX
XX Claim 2; SEQ ID NO 104; 53pp; English.
XX
CC The present invention relates to a method (M1) for diagnosing breast
CC cancer in a patient. The method comprises comparing the level of
CC expression of one or more genes or gene products in a biological sample
CC from the patient with that in a normal patient sample, where a difference
CC in the gene expression in the first sample compared to that in the second
CC sample is a diagnostic of the disease. Also claimed are: method (M2) for
CC distinguishing between normal and disease tissues; method (M3) for
CC monitoring the response of a breast cancer patient to treatment with an
CC anti-cancer agent; method (M4) for identifying a compound for treating
CC breast cancer; and an array for distinguishing between normal and disease
CC tissues comprising two or more probes corresponding to genes selected
CC from ADR98995-ADR99121 or comprising two or more polypeptides selected
CC from ADR99122-ADR99248. In M1 and M2 the genes are selected from ADR98995
CC -ADR99121 and the gene products are polypeptides selected from ADR99122-
CC ADR99248. M1 is useful for diagnosing breast cancer. M2 and the array are
CC useful for distinguishing between normal and disease tissue. M3 is useful
CC for monitoring the response of a breast cancer patient to treatment with
CC an anti-cancer agent. M4 is useful for identifying a compound for
CC treating breast cancer. Note: The sequence data for this patent did not
CC form part of the printed specification, but was obtained in electronic
CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 2243 BP; 661 A; 443 C; 525 G; 549 T; 0 U; 65 Other;
Query Match 79.2%; Score 19.8; DB 13; Length 2243;
Best Local Similarity 87.5%; Pred. No. 48;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 AAAAATCTCTGCTAGGCTTTTCAGT 24
Db 2095 AAAAATCTCTGCTAGGCTTTTCAGT 2072
RESULT 4
ABT22339/c
ID ABT22339 standard; DNA; 439 BP.
XX AC
XX ABT22339;
XX DT 16-APR-2003 (first entry)
```

```
XX Breast cancer marker gene SEQ ID No 712.
DE
XX Cytostatic; vaccine; breast cancer marker gene; breast mass; immunogen;
KW chemotherapy; tumour burden; bait protein; two-hybrid; three-hybrid;
KW surrogate marker gene; pharmacodynamic marker gene; transgenic animal;
KW human; ds.
XX
XX Homo sapiens.
XX
XX WO200285298-A2.
XX
XX 31-OCT-2002.
XX
XX 19-APR-2002; 2002WO-US012612.
XX
XX 20-APR-2001; 2001US-0285163P.
XX
XX (MILL-) MILLENNIUM PHARM INC.
XX
XX Lillie J, Palermo A, Wang Y, Steinmann K, Elias J, Mertens M;
XX
XX WPI; 2003-093053/08.
XX
XX Novel isolated polypeptide encoded by breast cancer marker gene, useful
PT for diagnosing, staging, monitoring, prognosing and treating diseases
PT associated with breast cancer.
XX
XX Disclosure; Page 189; 725pp; English.
XX
XX The invention relates to an isolated polypeptide encoded by a breast
CC cancer marker gene comprising any of 141/21-805 nucleotide sequences,
CC given in the specification. The methods of the invention are useful for
CC diagnosing patients having an identified breast mass or symptoms
CC associated with breast cancer, to diagnose breast cancer or its
CC precursors, and for monitoring the efficacy of treatment of a breast
CC cancer patient (e.g. efficacy of chemotherapy). The methods are also
CC useful for evaluating a patient before, after or during therapy, to
CC evaluate the reduction in a tumour burden. The breast cancer marker gene
CC proteins are useful as immunogens for raising antibodies, by immunising a
CC mammal with a breast cancer marker protein. The marker proteins are
CC useful as bait proteins in a two-hybrid or three-hybrid assay, to
CC identify other proteins which bind to or interact with the marker
CC proteins. The breast cancer marker genes are useful as surrogate marker
CC genes for one or more disorders, disease states or conditions leading to
CC disease states, in particular, breast cancers. The breast cancer marker
CC genes are useful as pharmacodynamic marker genes. An antibody which
CC selectively binds to a protein of a breast cancer marker gene is useful
CC for treating cancers, particularly breast cancers. The host cell of the
CC invention is useful for producing non-human transgenic animals. This
CC polynucleotide sequence represents one of the breast cancer marker genes
CC of the invention
XX
XX Sequence 439 BP; 111 A; 92 C; 98 G; 125 T; 0 U; 13 Other;
SQ
Query Match 76.8%; Score 19.2; DB 10; Length 439;
Best Local Similarity 87.5%; Pred. No. 71;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 AAAAACTCCTGTAGGCTTTTCAGT 24
Db 166 AAAAAATTCCTGTAGGTTTCATT 143
RESULT 5
ADU14171/c
ID ADU14171 standard; DNA; 545 BP.
XX
XX ADU14171;
XX
XX 27-JAN-2005 (first entry)
XX
XX Solid tumour prognosis gene seqid 4610.
DE
```

```
XX cytotstatic; gene therapy; expression profile; solid tumour;
KW peripheral blood mononuclear cell; PBMC; prognosis; ds.
XX
XX Unidentified.
XX
XX WO2004097052-A2.
XX
XX 11-NOV-2004.
XX
XX 29-APR-2004; 2004WO-US013587.
XX
XX 29-APR-2003; 2003US-0466067P.
XX
XX 23-JAN-2004; 2004US-0538246P.
XX
XX (AMHP ) WYETH.
XX
XX (STRA/) STRAHS A.
XX
XX Strahs A, Trepicchio WL, Burczynski ME, Twine NC, Slonim DK;
XX Immerman F, Dorner AJ;
XX
XX WPI; 2004-804779/79.
XX
XX A method, useful for prognosing and treating solid tumor, comprises
PT comparing an expression profile of a gene expressed in peripheral blood
PT mononuclear cells to a reference expression profile of a gene.
XX
XX Disclosure; Page; 111pp; English.
XX
XX The invention describes a method comprising comparing an expression
CC profile of at least one gene in a peripheral blood sample of a patient to
CC at least one reference expression profile of the at least one gene, where
CC the patient has a solid tumour, and each of the gene is differentially
CC expressed in peripheral blood mononuclear cells (PBMCs) of a first class
CC of patients as compared to PBMCs of a second class of patients, where
CC both the first and second classes is a subcluster formed by an
CC each of the first and second classes is a subcluster formed by an
CC unsupervised clustering analysis of gene expression profiles in PBMCs of
CC a population of patients who have the solid tumour, and where the
CC majority of the first class of patients has a first clinical outcome, and
CC the majority of the second class of patients has a second clinical
CC outcome. Also described are: a system comprising (i) a memory or a
CC storage medium including data that represent an expression profile of at
CC least one gene in a peripheral blood sample of a patient who has a solid
CC tumour, (ii) at least another storage medium including data that
CC represent at least one reference expression profile of the gene, (iii) a
CC program capable of comparing the expression profile to the reference
CC expression profile, and (iv) a processor capable of executing the
CC program, where expression levels of the gene in peripheral blood
CC mononuclear cells of patients who have the solid tumour correlate with
CC clinical outcomes of the patients; and a nucleic acid or protein array
CC comprising concentrated probes for solid tumour prognosis genes, where
CC each of the solid tumour prognosis genes is differentially expressed in
CC PBMCs of a first class of patients as compared to PBMCs of a second class
CC of patients, where both the first and second classes of patients have a
CC solid tumour, and where the first class of patients has a first clinical
CC outcome, and the second class of patients has a second clinical outcome.
CC The method, system, and array are useful for prognosing and treating
CC solid tumours. This sequence represents a solid tumour prognosis gene of
CC the invention. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_ftc_sequences.
XX
XX Sequence 545 BP; 145 A; 116 C; 97 G; 159 T; 0 U; 28 Other;
SQ
Query Match 76.8%; Score 19.2; DB 13; Length 545;
Best Local Similarity 87.5%; Pred. No. 73;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 AAAAACTCCTGTAGGCTTTTCAGT 24
Db 452 AAAAAATTCCTGTAGGTTTCATT 429
```

KW	coronary heart disease; ss.	
XX		
OS	Homo sapiens.	
XX		
PN	WO200157273-A2.	
XX		
PD	09-AUG-2001.	
XX		
XX	30-JAN-2001; 2001WO-US000664.	
PF		
XX	04-FEB-2000; 2000US-0180312P.	
PR	26-MAY-2000; 2000US-0207456P.	
PR	30-JUN-2000; 2000US-00608408.	
PR	03-AUG-2000; 2000US-00632366.	
PR	21-SEP-2000; 2000US-0234687P.	
PR	27-SEP-2000; 2000US-0236359P.	
PR	04-OCT-2000; 2000GB-00024263.	
XX		
XX	(MOLE-) MOLECULAR DYNAMICS INC.	
PA		
XX		
PI	Penn SG, Hanzel DK, Chen W, Rank DR;	
XX		
DR	WPI; 2001-48898/53.	
XX		
XX	Human genome-derived single exon nucleic acid probes useful for analyzing	
XX	gene expression in human adult liver.	
PT		
XX		
PS	Claim 1; SEQ ID NO 12702; 658pp; English.	
XX		
CC	The invention relates to a single exon nucleic acid probe (SENP) (I) for	
CC	measuring human gene expression in a sample derived from human adult	
CC	liver, comprising one of 13109 defined nucleotide sequences given in the	
CC	specification (or complements/ fragments). The probe hybridizes at high	
CC	stringency to a nucleic acid molecule expressed in the human adult liver.	
CC	(I) may be used for predicting, measuring and displaying gene expression	
CC	in samples derived from human adult liver. The genes identified may be	
CC	involved in genetic liver diseases such as cirrhosis,	
CC	hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is	
CC	associated with coronary heart disease. ABS25011-ABS51005 represent human	
CC	liver single exon nucleic acid probes of the invention. Note: The	
CC	sequence information for this patent does not appear in the printed	
CC	specification but was obtained in electronic format directly from WIPO at	
CC	ftp.wipo.int/pub/published_pct_sequences	
XX		
SQ	Sequence 562 BP; 202 A; 108 C; 133 G; 119 T; 0 U; 0 Other;	
	Query Match 76.8%; Score 19.2; DB 4; Length 562;	
	Best Local Similarity 87.5%; Pred. No. 74;	
	Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;	
QY	2 AAAAAGCTCTGTAGGCTTTGATG 25	
Db	29 AAAAAGCTCTGTATGCTTTGATTG 6	
RESULT 8		
ABS3712/c		
ID	ABS3712 standard; DNA; 562 BP.	
XX		
XX	ABS371206;	
AC		
XX		
XX	19-AUG-2002 (first entry)	
DT		
XX		
DE	Human genome-derived single exon probe from lung SEQ ID No 12097.	
XX		
XX	Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;	
KW	chronic obstructive pulmonary disease; interstitial lung disease;	
KW	familial idiopathic pulmonary fibrosis; neurofibromatosis;	
KW	tuberosus sclerosis; Gaucher's disease; Niemann-Pick disease;	
KW	Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;	
KW	pulmonary histiocytosis; lymphangioleiomyomatosis; Karagenar syndrome;	
KW	pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;	
KW	primary ciliary dyskinesia; pulmonary hypertension;	

KW	coronary heart disease; ss.	
XX		
OS	Homo sapiens.	
XX		
PN	WO200157273-A2.	
XX		
PD	09-AUG-2001.	
XX		
XX	30-JAN-2001; 2001WO-US000664.	
PF		
XX	04-FEB-2000; 2000US-0180312P.	
PR	26-MAY-2000; 2000US-0207456P.	
PR	30-JUN-2000; 2000US-00608408.	
PR	03-AUG-2000; 2000US-00632366.	
PR	21-SEP-2000; 2000US-0234687P.	
PR	27-SEP-2000; 2000US-0236359P.	
PR	04-OCT-2000; 2000GB-00024263.	
XX		
XX	(MOLE-) MOLECULAR DYNAMICS INC.	
PA		
XX		
PI	Penn SG, Hanzel DK, Chen W, Rank DR;	
XX		
DR	WPI; 2001-483446/52.	
XX		
XX	Single exon nucleic acid probes for analyzing gene expression in human	
XX	brains.	
PT		
XX		
PS	Example 4; SEQ ID NO 12369; 650pp + Sequence Listing; English.	
XX		
CC	The present invention provides a number of single exon nucleic acid	
CC	probes which are derived from genomic sequences expressed in the human	
CC	brain. They can be used to measure gene expression in brain cell samples,	
CC	which may enable the diagnosis and improved treatment of nervous system	
CC	diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,	
CC	epilepsy and cancers. The present sequence is one of the probes of the	
CC	invention	
XX		
SQ	Sequence 562 BP; 202 A; 108 C; 133 G; 119 T; 0 U; 0 Other;	
	Query Match 76.8%; Score 19.2; DB 4; Length 562;	
	Best Local Similarity 87.5%; Pred. No. 74;	
	Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;	
QY	2 AAAAAGCTCTGTAGGCTTTGATG 25	
Db	29 AAAAAGCTCTGTATGCTTTGATTG 6	
RESULT 7		
ABS37712/c		
ID	ABS37712 standard; DNA; 562 BP.	
XX		
XX	ABS37712;	
AC		
XX		
DT	25-FEB-2003 (first entry)	
XX		
XX	Human liver single exon probe, SEQ ID No 12702.	
DE		
XX		
KW	Human; single exon nucleic acid probe; liver; cirrhosis;	
KW	hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;	

DE Human breast cancer antigen polynucleotide seq id 6.
KW breast cancer; breast cancer diagnosis; breast cancer antigen; gene; ss.
XX
OS Homo sapiens.
PN US2003108888-A1.
XX
PD 12-JUN-2003.
XX
PF 15-MAY-2002; 2002US-00146473.
XX
PR 15-MAY-2001; 2001US-0291150P.
XX
PA (LUDW-) LUDWIG INST CANCER RES.
XX
PI Scanlan MJ, Gout I, Stockert E, Old LJ, Gure A, Chen Y;
XX
DR WPI; 2003-829397/77.
DR P-PSDB; ADC35080.
XX
PT Diagnosing breast cancer in subject by obtaining biological sample from
PT subject, contacting sample with breast cancer-associated polypeptides,
PT determining specific binding between polypeptides and agents in sample.
XX
PS Claim 1; SEQ ID NO 6; 173pp; English.
XX
CC The invention describes a method of diagnosing breast cancer in subject
CC comprising contacting biological sample from subject with at least two
CC different breast cancer-associated polypeptides (I) encoded by nucleic
CC acid molecules (II) comprising sequence chosen from 42 fully defined
CC sequences as given in specification, determining specific binding between
CC (I) and agents in sample, where presence of the binding is diagnostic for
CC breast cancer. The method is useful for diagnosing breast cancer in a
CC subject. The sample is blood, lymph node fluid or breast discharge fluid.
CC This sequence encodes a breast cancer antigen.
XX
SQ Sequence 2587 BP; 711 A; 608 C; 670 G; 598 T; 0 U; 0 Other;

Query Match 76.8%; Score 19.2; DB 10; Length 2587;
Best Local Similarity 87.5%; Pred. No. 93;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAACTCTGTAGGCTTTCAGT 24
DB 2463 AAAAAATTCCTGTAGGTTTTCATT 2440

RESULT 11
ADN03792/C
ID ADN03792 standard; cDNA; 2587 BP.
XX
AC ADN03792;
XX
DT 01-JUL-2004 (first entry)
XX
DE Antipsoriatic cDNA sequence #94.
XX
KW ds; gene; antipsoriatic; gene therapy; psoriasis; diagnosis.
XX
OS Homo sapiens.
XX
FN WO2004028479-A2.
XX
PD 08-APR-2004.
XX
PF 25-SEP-2003; 2003WO-US030907.
XX
PR 25-SEP-2002; 2002US-0414006P.
XX
PA (GETH) GENENTECH INC.
XX
PI Bodary S, Clark H, Jackman J, Schoenfeld J, Williams PM, Wood WT;

PI Wu TD;
XX
DR WPI; 2004-305105/28.
DR P-PSDB; ADN03793.
XX
XX New PRO nucleic acid or polypeptide, useful for preparing a
PT pharmaceutical composition for diagnosing or treating psoriasis in a
PT mammal.
XX
PS Claim 1; SEQ ID NO 186; 3069pp; English.
XX
CC The invention relates to novel polynucleotide and polypeptides for
CC treating psoriasis or a sequence having at least 80% identity to the
CC above sequences. The nucleic acid is useful for preparing a composition
CC for diagnosing or treating psoriasis in a mammal. This sequence
CC corresponds to one of the polynucleotides of the invention.
XX
SQ Sequence 2587 BP; 711 A; 608 C; 670 G; 598 T; 0 U; 0 Other;

Query Match 76.8%; Score 19.2; DB 12; Length 2587;
Best Local Similarity 87.5%; Pred. No. 93;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAACTCTGTAGGCTTTCAGT 24
DB 2463 AAAAAATTCCTGTAGGTTTTCATT 2440

RESULT 12
ADP13389/C
ID ADP13389 standard; DNA; 2587 BP.
XX
AC ADP13389;
XX
DT 26-AUG-2004 (first entry)
XX
DE Renal cell carcinoma differentially expressed gene #125.
XX
KW ds; diagnosis; non-blood disease; solid tumor; gene expression;
KW peripheral blood mononuclear cell; renal cell carcinoma; prostate cancer;
KW head/neck cancer; differential expression.
XX
OS Homo sapiens.
XX
FN WO2004048933-A2.
XX
PD 10-JUN-2004.
XX
PF 21-NOV-2003; 2003WO-US037481.
XX
PR 21-NOV-2002; 2002US-0427982P.
PR 03-APR-2003; 2003US-0459782P.
XX
PA (AMHP) WYETH.
PA (TWIN/) TWINE N C.
PA (BURC/) BURCZYNSKI M E.
PA (TREP/) TREPICCHIO W L.
PA (DORN/) DORNER A.
PA (STOV/) STOVER J A.
PA (SLON/) SLONI D K.
XX
PI Twine NC, Burczynski ME, Trepicchio WL, Dorner A, Stover JA;
PI Sloni DK;
XX
DR WPI; 2004-460799/43.
XX
PT Diagnosing non-blood disease such as solid tumor, involves comparing
PT differential expression profile of specific genes in peripheral blood
PT sample of subject with reference expression profile of specific genes.
XX
PS Disclosure; SEQ ID NO 125; 350pp; English.
XX
CC The invention relate to a method of diagnosing (M1) non-blood disease

CC such as solid tumor by providing peripheral blood sample of human having
 CC non-blood disease, and comparing an expression profile of specific genes
 CC in the peripheral blood sample to reference expression profile of the
 CC genes, where each of the genes is differentially expressed in peripheral
 CC blood mononuclear cells (PBMCs) of patients having the disease as
 CC compared to PBMCs of normal humans. The method is useful for diagnosing
 CC non-blood disease such as solid tumor. The solid tumor is chosen from
 CC renal cell carcinoma (RCC), prostate cancer and head/neck cancer. The
 CC peripheral blood sample comprises enriched PBMCs. The peripheral blood
 CC sample is a whole blood sample (claimed). (M1) is useful for identifying
 CC genes that are differentially expressed in peripheral blood samples
 CC isolated at different stages of progression, development or treatment of
 CC RCC and/or other solid tumors. This sequence corresponds to a gene that
 CC is differentially expressed and detected by the method of the invention.
 CC (Note: this sequence is not given as part of the printed specification
 CC but was obtained from WIPO in electronic format at
 CC ftp.wipo./pub/published_pct_sequences).
 XX
 SQ Sequence 2587 BP; 711 A; 608 C; 670 G; 598 T; 0 U; 0 Other;

Query Match 76.8%; Score 19.2; DB 12; Length 2587;
 Best Local Similarity 87.5%; Pred. No. 93;
 Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAACTCTGTAGGCTTTTCAGT 24
 ||||| ||||| ||||| ||||| |||||
 Db 2463 AAAAAATTCCTGTAGGTTTCATT 2440

RESULT 13
 ADP54303/c
 ID ADP54303 standard; cDNA; 2587 BP.
 XX
 AC ADP54303;
 XX
 DT 18-NOV-2004 (first entry)
 XX
 DE Human PRO cDNA sequence SEQ ID NO:279.

XX human; PRO; immune related disease; inflammatory immune response;
 KW immune response stimulation; anti-allergic; antianaemic; antiarthritic;
 KW antiasthmatic; antidiabetic; antiinflammatory; antipsoriatic;
 KW antirheumatic; antithyroid; CNS; dermatological; gastrointestinal;
 KW haemostatic; hepatotropic; immunostimulant; immunosuppressive; muscular;
 KW nephrotropic; neuroprotective; osteopathic; respiratory; vasotropic;
 KW virucide; gene therapy; gene; ss.

XX Homo sapiens.
 XX
 XX WO2004039956-A2.
 XX
 XX 13-MAY-2004.
 XX
 XX 28-OCT-2003; 2003WO-US034381.
 XX
 XX 29-OCT-2002; 2002US-0422472P.
 XX
 XX (GETH) GENENTECH INC.
 XX
 XX Aggarwal S, Clark H, Gurney AL, Schoenfeld J, Williams PM;
 PI Wood WI, Wu TD;
 XX
 XX WPI; 2004-376182/35.
 DR P-PSDB; ADP54304.
 XX

XX New PRO polynucleotides and polypeptides, useful in diagnosing
 PT and treating an immune related disease, e.g. systemic lupus
 PT erythematosus, rheumatoid arthritis, diabetes mellitus or asthma and in
 PT stimulating an immune response.

XX Claim 2; SEQ ID NO 279; 3009pp; English.

XX The present invention describes an isolated PRO nucleic acid (I). Also

CC described: (1) a vector comprising (1); (2) a host cell comprising the
 CC vector of (1); (3) a process for producing a PRO polypeptides; (4) an
 CC isolated PRO polypeptide; (5) a chimeric molecule comprising the
 CC polypeptide of (4) fused to a heterologous amino acid sequence; (6) an
 CC antibody which specifically binds to a polypeptide of (4); (7) a
 CC composition of matter comprising a polypeptide of (4), an agonist or
 CC antagonist of the polypeptide or an antibody that binds to the
 CC polypeptide in combination with a carrier; (8) an article of manufacture
 CC comprising a container, a label on the container and a composition of
 CC matter of (7); (9) a method of treating an immune related disease in a
 CC mammal; (10) a method for determining the presence of a PRO polypeptide
 CC in a sample suspected of having the polypeptide; (11) a method of
 CC diagnosing an immune related disease or an inflammatory immune response
 CC in mammal; (12) a method of identifying a compound that inhibits or
 CC mimics the activity of or expression of a gene encoding a PRO polypeptide
 CC ; and (13) a method of stimulating the immune response in a mammal. The
 CC PRO sequences have anti-allergic, antianaemic, antiarthritic,
 CC antiasthmatic, antidiabetic, antiinflammatory, antipsoriatic,
 CC antirheumatic, antithyroid, CNS, dermatological, immunosuppressive, muscular,
 CC haemostatic, hepatotropic, immunostimulant, immunosuppressive, osteopathic,
 CC nephrotropic, neuroprotective, osteopathic, respiratory, vasotropic and
 CC virucide activities, and can be used in gene therapy. The nucleic acid
 CC (I) and the encoded polypeptides, compositions, kits and methods are
 CC useful in diagnosing and treating an immune related disease and in
 CC stimulating an immune response. The present sequence represents a human
 CC PRO nucleotide sequence from the present invention.

XX
 SQ Sequence 2587 BP; 711 A; 608 C; 670 G; 598 T; 0 U; 0 Other;

Query Match 76.8%; Score 19.2; DB 13; Length 2587;
 Best Local Similarity 87.5%; Pred. No. 93;
 Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAACTCTGTAGGCTTTTCAGT 24
 ||||| ||||| ||||| ||||| |||||
 Db 2463 AAAAAATTCCTGTAGGTTTCATT 2440

RESULT 14
 ADP24764/c
 ID ADP24764 standard; cDNA; 2587 BP.
 XX
 AC ADP24764;
 XX
 DT 18-NOV-2004 (first entry)
 XX
 DE PRO polypeptide encoding cDNA SEQ ID NO:1942.
 XX
 XX ss; gene; PRO; antiinflammatory; antiarthritic; antirheumatic;
 KW immunosuppressive; osteopathic; antidiabetic; dermatological;
 KW antipsoriatic; anti-allergic; antiasthmatic; hepatotropic; respiratory;
 KW gene therapy; immune system.
 XX
 XX Unidentified.
 XX
 XX WO2004041170-A2.
 XX
 XX 21-MAY-2004.
 XX
 XX 30-OCT-2003; 2003WO-US034312.
 XX
 XX 01-NOV-2002; 2002US-0423394P.
 XX
 XX (GETH) GENENTECH INC.

XX Clark H, Schoenfeld J, Van Lookeren M, Williams PM, Wood WI;
 PI Wu TD;
 XX
 XX WPI; 2004-419628/39.
 DR P-PSDB; ADP24765.

XX New PRO polypeptides and polynucleotides, useful for treating e.g.
 PT erythematosus, rheumatoid arthritis, diabetes mellitus, immune-mediated

PT renal disease, or demyelinating diseases of the central or peripheral
PT nervous system.
XX
PS Claim 1; SEQ ID NO 1942; 2940pp; English.
XX
CC The invention relates to a novel isolated nucleic acid and the PRO
CC polypeptide encoded by it. A protein of the invention has
CC antiinflammatory, antiarthritic, antirheumatic, immunosuppressive,
CC osteopathic, antidiabetic, dermatological, antipsoriatic, antiallergic,
CC antiasthmatic, hepatotropic, and respiratory activity. A polynucleotide
CC of the invention may have a use in gene therapy. The PRO polypeptide, its
CC agonist, antagonist, or antibody that specifically binds to the
CC polypeptide is useful for treating an immune related disorder such as
CC systemic lupus erythematosus, rheumatoid arthritis, osteoarthritis,
CC juvenile chronic arthritis, a spondyloarthropathy, systemic sclerosis, an
CC idiopathic inflammatory myopathy, Sjogren's syndrome, systemic
CC vasculitis, sarcoidosis, autoimmune haemolytic anaemia, autoimmune
CC thrombocytopenia, thyroiditis, diabetes mellitus, immune-mediated renal
CC disease, a demyelinating disease of the central or peripheral nervous
CC system, idiopathic demyelinating polyneuropathy, Guillain-Barre syndrome,
CC a chronic inflammatory demyelinating polyneuropathy, a hepatobiliary
CC disease, infectious or autoimmune chronic active hepatitis, primary
CC biliary cirrhosis, granulomatous hepatitis, sclerosing cholangitis,
CC inflammatory bowel disease, gluten-sensitive enteropathy, Whipple's
CC disease, an autoimmune or immune-mediated skin disease, a bullous skin
CC disease, erythema multiforme, contact dermatitis, psoriasis, an allergic
CC disease, asthma, allergic rhinitis, atopic dermatitis, food
CC hypersensitivity, urticaria, an immunologic disease of the lung,
CC eosinophilic pneumonia, idiopathic pulmonary fibrosis, hypersensitivity
CC pneumonitis, a transplantation associated disease, graft rejection or
CC graft-versus-host disease. The present sequence encodes a PRO protein of
CC the invention.
XX
SQ Sequence 2587 BP; 711 A; 608 C; 670 G; 598 T; 0 U; 0 Other;

Query Match 76.8%; Score 19.2; DB 13; Length 2587;
Best Local Similarity 87.5%; Pred. No. 93;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 AAAAAAATCCTGTAGGTTTCATT 24
Db 2463 AAAAAAATCCTGTAGGTTTCATT 2440

RESULT 15
ADU06152/c
ID ADU06152 standard; DNA; 2587 BP.
AC ADU06152;
XX
DT 27-JAN-2005 (first entry)
XX
DE Novel bronchial cancer-associated human gene SeqID376.
XX
KW bronchial' cancer; cytostatic; tumour-associated protein;
KW cancer detection; metastasis; tumour; gene; ds; human.
XX
OS Homo sapiens.
XX
PN DE10316701-A1.
XX
PD 04-NOV-2004.
XX
PF 09-APR-2003; 2003DE-01016701.
XX
PR 09-APR-2003; 2003DE-01016701.
XX
PA (HINZ/) HINZMANN B.
PA (HERM/) HERMANN K.
PA (CAST/) HEIDEN CASTANOS-VELEZ E.
XX
PI Mennerich D, Bruemendorf T, Heiden E, Hermann K, Kinneemann H;
PI Li X, Roepcke S, Staub E, Hinzmann B, Rosenthal A, Pilarsky C;

XX WPI; 2004-786403/78.
DR P-PSDB; ADU06639.
XX
PT New nucleic acid, and derived proteins, useful for diagnosis of bronchial
PT cancer and in screening for therapeutic and diagnostic agents.
XX
PS Claim 1; SEQ ID NO 376; 1381pp; German.
XX
CC This invention relates to a novel isolated nucleic acid associated with
CC bronchial cancer comprising 489 defined sequences given in the
CC specification. The invention may be useful for the production of
CC compounds with a cytostatic activity through the inhibition of expression
CC or activity of tumour-associated proteins. The novel DNA sequences and
CC the proteins/peptides encoded by them are used for detecting bronchial
CC cancer or determining the risk of developing it and to screen for
CC specific binding partners of the DNA or protein sequences, where the
CC binding partners are potentially useful as agents for treating or
CC diagnosing bronchial cancer. The DNA or protein sequences can also be
CC used for prognosis, detection of metastases and for secondary treatment
CC (of tumours that have been stabilised or are no longer detectable).
CC Detecting abnormal expression of the DNA sequences provides early
CC diagnosis of bronchial cancers. The present sequence is that of a novel
CC bronchial cancer-associated human gene sequence of the invention.
XX
SQ Sequence 2587 BP; 711 A; 608 C; 670 G; 598 T; 0 U; 0 Other;
Query Match 76.8%; Score 19.2; DB 13; Length 2587;
Best Local Similarity 87.5%; Pred. No. 93;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 AAAAAAATCCTGTAGGTTTCATT 24
Db 2463 AAAAAAATCCTGTAGGTTTCATT 2440

RESULT 16
ADY20501/c
ID ADY20501 standard; DNA; 2587 BP.
XX
AC ADY20501;
XX
DT 05-MAY-2005 (first entry)
XX
DE DNA encoding a PRO polypeptide, SEQ ID NO 6307.
XX
KW Antinflammatory; Immune disorder; Dermatological; Immunosuppressive;
KW Antirheumatic; Antiarthritic; Osteopathic; Hemostatic; Antianemic;
KW Antithyroid; Antidiabetic; Nephrotropic; CNS-Gen.; Hepatotropic;
KW Virucide; Gastrointestinal-Gen.; Antipsoriatic; Antiasthmatic;
KW Antiallergic; ds; gene; diagnosis.
XX
XX Homo sapiens.
XX
PN WO2005016962-A2.
XX
PD 24-FEB-2005.
XX
PF 11-AUG-2004; 2004WO-US026249.
XX
PR 11-AUG-2003; 2003US-0493546P.
XX
PA (GETH) GENENTECH INC.
XX
PI Abbas A, Clark H, Ouyang W, Williams MP, Wood WI, Wu TD;
XX
DR WPI; 2005-182330/19.
XX
PT New nucleic acid encoding PRO polypeptide, useful for diagnosing and
PT treating an immune related disorder, e.g. systemic lupus erythematosus,
PT rheumatoid arthritis, osteoarthritis, thyroiditis, or diabetes mellitus.
XX
PS Claim 1; SEQ ID NO 6307; 158pp; English.

XX The invention relates to an isolated nucleic acid encoding a PRO
CC polypeptide. The polypeptide, agonist or an antagonist, antibody,
CC composition, and method are useful for diagnosing and treating an immune
CC related disorder, e.g. systemic lupus erythematosus, rheumatoid
CC arthritis. The present sequence represents a PRO polypeptide.
XX
SQ Sequence 2587 BP; 711 A; 608 C; 670 G; 598 T; 0 U; 0 Other;

Query Match 76.8%; Score 19.2; DB 14; Length 2587;
Best Local Similarity 87.5%; Pred. No. 93;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAACTCCTGTAGGCTTTTCAGT 24
|||||
Db 2463 AAAAAATTCCTGTAGGTTTTCATT 2440
|||||

RESULT 17
ADY14680/c
ID ADY14680 standard; DNA; 2587 BP.
XX
AC ADY14680;
XX
DT 05-MAY-2005 (first entry)
XX
DE DNA encoding a PRO polypeptide, SEQ ID NO 486.
XX
KW Antinflammatory; Immune disorder; Dermatological; Immunosuppressive;
KW Antirheumatic; Antiarthritic; Osteopathic; Hemostatic; Antianemic;
KW Anticholesteric; Antidiabetic; Nephrotropic; CNS-Gen.; Hepatotropic;
KW Virucide; Gastrointestinal-Gen.; Antipsoriatic; Antiasthmatic;
KW Antiallergic; ds; gene; diagnosis.
XX
OS Homo sapiens.
XX
PN WO2005016962-A2.
XX
PD 24-FEB-2005.
XX
PF 11-AUG-2004; 2004WO-US026249.
XX
PR 11-AUG-2003; 2003US-0493546P.
XX
PA (GETH) GENENTECH INC.
XX
PI Abbas A, Clark H, Ouyang W, Williams MP, Wood WI, Wu TD;
XX
DR WPI; 2005-182330/19.
XX
PT New nucleic acid encoding PRO polypeptide, useful for diagnosing and
PT treating an immune related disorder, e.g. systemic lupus erythematosus,
PT rheumatoid arthritis, osteoarthritis, thyroiditis, or diabetes mellitus.
XX
PS Claim 1; SEQ ID NO 486; 158pp; English.
XX
CC The invention relates to an isolated nucleic acid encoding a PRO
CC polypeptide. The polypeptide, agonist or an antagonist, antibody,
CC composition, and method are useful for diagnosing and treating an immune
CC related disorder, e.g. systemic lupus erythematosus, rheumatoid
CC arthritis. The present sequence represents a DNA encoding a PRO
CC polypeptide.
XX
SQ Sequence 2587 BP; 711 A; 608 C; 670 G; 598 T; 0 U; 0 Other;

Query Match 76.8%; Score 19.2; DB 14; Length 2587;
Best Local Similarity 87.5%; Pred. No. 93;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAACTCCTGTAGGCTTTTCAGT 24
|||||
Db 2463 AAAAAATTCCTGTAGGTTTTCATT 2440
|||||

RESULT 18
ADE53710/c
ID ADE53710 standard; cDNA; 3886 BP.
XX
AC ADE53710;
XX
DT 29-JAN-2004 (first entry)
XX
DE Human prostate cancer cDNA #57.
XX
KW Human; prostate cancer; ss; cDNA combination; differential expression;
KW gene.
XX
OS Homo sapiens.
XX
PN US2003190640-A1.
XX
PD 09-OCT-2003.
XX
PF 29-MAY-2002; 2002US-00252157.
XX
PR 31-MAY-2001; 2001US-0295048P.
XX
PA (FARI/) FARIS M.
PA (PEAR/) PEARSON C I.
XX
PI Faris M, Pearson CI;
XX
DR WPI; 2003-831619/77.
XX
PT New combination comprising cDNAs that are differentially expressed in
PT prostate cancer, useful for diagnosing, treating or monitoring the
PT progression of treatment of prostate cancer.
XX
PS Claim 1; SEQ ID NO 57; 42pp; English.
XX
CC The invention relates to a combination comprising a number of cDNAs
CC expressed in prostate cancer. The invention also relates to a method for
CC detecting differential expression of one or more cDNAs in a sample
CC containing nucleic acids by hybridising a substrate with the nucleic
CC acids, thus forming one or more hybridisation complexes, detecting
CC hybridisation complex formation and comparing the complexes formed with
CC standard complexes, where differences between the standard and the sample
CC complex formation indicate differential expression of cDNAs in the
CC sample. The differential expression is diagnostic of prostate cancer. The
CC invention also relates to proteins and antibodies related to the cDNAs.
CC The combination is useful for diagnosing, treating or monitoring the
CC progression of treatment of prostate cancer. The antibodies are useful
CC for detecting prostate cancer. This sequence represents a human prostate
CC cancer cDNA of the invention.
XX
SQ Sequence 3886 BP; 1091 A; 775 C; 819 G; 1030 T; 0 U; 171 Other;

Query Match 76.8%; Score 19.2; DB 10; Length 3886;
Best Local Similarity 87.5%; Pred. No. 99;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAACTCCTGTAGGCTTTTCAGT 24
|||||
Db 2445 AAAAAATTCCTGTAGGTTTTCATT 2422
|||||

RESULT 19
ADL61917/c
ID ADL61917 standard; DNA; 5569 BP.
XX
AC ADL61917;
XX
DT 20-MAY-2004 (first entry)
XX
DE Human ovarian cancer DNA marker #20129.
XX

KW Human; ovarian cancer; ds; tumour; cytostatic; DNA marker.

XX Homo sapiens.

XX WO200170979-A2.

XX 27-SEP-2001.

XX 21-MAR-2001; 2001WO-US009126.

XX 21-MAR-2000; 2000US-0191031P.

XX 25-MAY-2000; 2000US-0207124P.

XX 15-JUN-2000; 2000US-0211940P.

XX 07-JUL-2000; 2000US-0216920P.

XX 25-JUL-2000; 2000US-0220661P.

XX 21-DEC-2000; 2000US-0257672P.

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX Lee J, Lillie J;

XX WPI; 2001-611502/70.

XX Novel isolated nucleic acid molecules (markers) overexpressed in ovarian

XX cancer cells as compared to their normal non-cancerous ovarian cells are

XX used to characterize stage, grade, histological type of ovarian cancer.

XX Disclosure; SEQ ID NO 20129; 106pp; English.

XX The invention relates to nucleic acid markers which are overexpressed in

XX ovarian cancer cells as compared to their expression in normal (i.e. non-

XX cancerous) ovarian cells. The invention also relates to polypeptides

XX encoded by the markers, antibodies that selectively bind to the

XX polypeptides, a method of inhibiting ovarian cancer in a patient at risk

XX of developing ovarian cancer involving inhibiting expression of a gene

XX corresponding to a marker of the invention and a method of treating a

XX patient afflicted with ovarian cancer comprising providing to cells of

XX the patient an antisense oligonucleotide complementary to a marker of the

XX invention. The markers are useful for assessing if a patient is afflicted

XX with ovarian cancer, which involves comparing the level of expression of

XX a marker in a patient sample and a normal level of expression of the

XX expression levels indicates ovarian cancer. The level of expression of a

XX marker corresponds to a secreted protein or to a transcribed

XX polynucleotide or its portion. The level of expression of the marker is

XX assessed by detecting the presence in the sample, a protein or protein

XX fragment corresponding to the marker. The presence of protein or protein

XX fragment is detected using an antibody that specifically binds with the

XX protein or protein fragment. Alternatively, the level of expression of

XX the marker is assessed by detecting the presence of a transcribed

XX polynucleotide which anneals with the marker or anneals with a portion of

XX the polynucleotide comprising the marker, under stringent conditions. The

XX marker is also used for monitoring the progression of ovarian cancer in a

XX patient which involves detecting expression of the marker in a patient

XX sample at a first point in time, repeating the method at a subsequent

XX time and comparing the level of expression. The method is carried out

XX using an ovarian tissue sample. A composition comprising a marker, cancer,

RESULT 20

ABV24840/c

ID ABV24840 standard; cDNA; 5570 BP.

XX AC ABV24840;

XX 16-SEP-2002 (first entry)

XX Human prostate expression marker cDNA 24831.

XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;

XX pharmacogenomic marker; gene; ss.

XX Homo sapiens.

XX WO200160860-A2.

XX 23-AUG-2001.

XX 20-FEB-2001; 2001WO-US005171.

XX 17-FEB-2000; 2000US-0183319P.

XX 16-MAR-2000; 2000US-0189862P.

XX 25-MAY-2000; 2000US-0207454P.

XX 09-JUN-2000; 2000US-0211314P.

XX 18-JUL-2000; 2000US-0219007P.

XX 13-DEC-2000; 2000US-0255281P.

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX Schlegel R, Endege WO, Monahan JE;

XX WPI; 2001-662795/76.

XX Novel isolated nucleic acid molecule associated with cancerous state of

XX prostate cells and correlating with presence of prostate cancer, useful

XX for detecting presence of prostate cancer, stage of prostate cancer.

XX Claim 1; Page 4764-4766; 11750pp; English.

XX The invention relates to an isolated nucleic acid molecule (I) comprising

XX a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the

XX specification or its complement. (I) is useful for: (a) assessing whether

XX a patient is afflicted with prostate cancer; (b) monitoring the

XX progression of prostate cancer in a patient; (c) assessing the efficacy

XX of a test compound to inhibit prostate cancer in a patient; (d) assessing

XX the efficacy of a therapy for inhibiting prostate cancer in a patient;

XX (e) selecting a composition for inhibiting prostate cancer in a patient;

XX (f) assessing the prostate cell carcinogenic potential of a compound; (g)

XX determining whether prostate cancer has metastasized in a patient; (h)

XX assessing the aggressiveness or indolence of prostate cancer in a patient

XX ; (i) is also useful as a pharmacodynamic or pharmacogenomic marker

XX Sequence 5570 BP; 1655 A; 1123 C; 1062 G; 1721 T; 0 U; 8 Other;

XX Query Match 76.8%; Score 19.2; DB 5; Length 5570;

XX Best Local Similarity 87.5%; Pred. No. 1.1e+02;

XX Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAACTCTGTAGGCTTTTCAGT 24

||||| ||||||| ||||||| |||||||

Db 317 AAAAAATTCCTGTAGGCTTTTCATT 294

||||| ||||||| ||||||| |||||||

KW Human; ovarian cancer; ds; tumour; cytostatic; DNA marker.

XX Homo sapiens.

XX WO200170979-A2.

XX 27-SEP-2001.

XX 21-MAR-2001; 2001WO-US009126.

XX 21-MAR-2000; 2000US-0191031P.

XX 25-MAY-2000; 2000US-0207124P.

XX 15-JUN-2000; 2000US-0211940P.

XX 07-JUL-2000; 2000US-0216920P.

XX 25-JUL-2000; 2000US-0220661P.

XX 21-DEC-2000; 2000US-0257672P.

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX Lee J, Lillie J;

XX WPI; 2001-611502/70.

XX Novel isolated nucleic acid molecules (markers) overexpressed in ovarian

XX cancer cells as compared to their normal non-cancerous ovarian cells are

XX used to characterize stage, grade, histological type of ovarian cancer.

XX Disclosure; SEQ ID NO 20129; 106pp; English.

XX The invention relates to nucleic acid markers which are overexpressed in

XX ovarian cancer cells as compared to their expression in normal (i.e. non-

XX cancerous) ovarian cells. The invention also relates to polypeptides

XX encoded by the markers, antibodies that selectively bind to the

XX polypeptides, a method of inhibiting ovarian cancer in a patient at risk

XX of developing ovarian cancer involving inhibiting expression of a gene

XX corresponding to a marker of the invention and a method of treating a

XX patient afflicted with ovarian cancer comprising providing to cells of

XX the patient an antisense oligonucleotide complementary to a marker of the

XX invention. The markers are useful for assessing if a patient is afflicted

XX with ovarian cancer, which involves comparing the level of expression of

XX a marker in a patient sample and a normal level of expression of the

XX expression levels indicates ovarian cancer. The level of expression of a

XX marker corresponds to a secreted protein or to a transcribed

XX polynucleotide or its portion. The level of expression of the marker is

XX assessed by detecting the presence in the sample, a protein or protein

XX fragment corresponding to the marker. The presence of protein or protein

XX fragment is detected using an antibody that specifically binds with the

XX protein or protein fragment. Alternatively, the level of expression of

XX the marker is assessed by detecting the presence of a transcribed

XX polynucleotide which anneals with the marker or anneals with a portion of

XX the polynucleotide comprising the marker, under stringent conditions. The

XX marker is also used for monitoring the progression of ovarian cancer in a

XX patient which involves detecting expression of the marker in a patient

XX sample at a first point in time, repeating the method at a subsequent

XX time and comparing the level of expression. The method is carried out

XX using an ovarian tissue sample. A composition comprising a marker, cancer,

RESULT 21

AAK84729/c

ID AAK84729 standard; DNA; 48037 BP.

XX AC AAK84729;

XX 07-NOV-2001 (first entry)

XX

DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:39541.
XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX Homo sapiens.
OS
XX
XX
PN WO200157182-A2.
XX
XX
XX 09-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US001354.
XX
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205513P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 08-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.

PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 12-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.

```
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-483426/52.
XX
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
XX useful for preventing, diagnosing and/or treating cancers and metastasis.
XX
XX Disclosure; SEQ ID NO 39541; 3071pp + Sequence Listing; English.
XX
XX AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
XX amino acid sequences given in AAK62170 to AAK61921. (I) have cytostatic
XX activity, and can be used in gene therapy and vaccine production. (I)
XX proteins and polynucleotides may be used in the prevention, diagnosis and
XX treatment of diseases associated with inappropriate (I) expression. For
XX example, they may be used to treat disorders associated with decreased
XX expression by rectifying mutations or deletions in a patient's genome
XX that affect the activity of (I) by expressing inactive proteins or to
XX supplement the patient's own production of (I). Additionally, (I)
XX polynucleotides may be used to produce the secreted (I), by inserting the
XX nucleic acids into a host cell and culturing the cell to express the
XX protein. (I) proteins and polynucleotides may be used to prevent,
XX diagnose and treat immune/haematopoietic-related diseases, especially
XX cancers and cancer metastases of haematopoietic-derived cells. AAK64703
XX to AAK87694 represent human immune/haematopoietic antigen genomic
XX sequences from the present invention. AAK54942 to AAK54950 and AAK82169
XX represent sequences used in the exemplification of the present invention
XX
XX SQ Sequence 48037 BP; 12892 A; 9814 C; 10466 G; 14865 T; 0 U; 0 Other;

Query Match 76.8%; Score 19.2; DB 4; Length 48037;
Best Local Similarity 87.5%; Pred. No. 1.5e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAAATCTCTGAGGTTTCATT 43132
Db 43155 AAAAAAATCTCTGAGGTTTCATT 43132

RESULT 22
AAK85983/c
ID AAK85983 standard; DNA; 48037 BP.
AC AAK85983;
XX
XX 07-NOV-2001 (first entry)
XX
XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:40795.
XX
XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
XX cytostatic; gene therapy; vaccine; metastasis; ds.
XX
XX Homo sapiens.
XX
XX W0200157182-A2.
XX
XX 09-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US001354.
XX
XX 31-JAN-2000; 2000US-0179065P.
XX 04-FEB-2000; 2000US-0180628P.
XX 24-FEB-2000; 2000US-0184664P.
XX 02-MAR-2000; 2000US-0186350P.
XX 16-MAR-2000; 2000US-0189874P.
XX 17-MAR-2000; 2000US-0190076P.
XX 18-APR-2000; 2000US-0198123P.
XX 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226688P.
PR 23-AUG-2000; 2000US-0227182P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
```

PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244611P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-483426/52.
XX
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and metastasis.
PT
XX Disclosure; SEQ ID NO 40795; 3071pp + Sequence Listing; English.
PS
XX AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
CC amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased

CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting the
CC nucleic acids into a host cell and culturing the cell to prevent, the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and cancer metastases of hematopoietic-related diseases, especially
CC to AAK87694 represent human immune/hematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
CC represent sequences used in the exemplification of the present invention
XX
SQ Sequence 48037 BP; 12892 A; 9814 C; 10466 G; 14865 T; 0 U; 0 Other;
Query Match 76.8%; Score 19.2; DB 4; Length 48037;
Best Local Similarity 87.5%; Pred. No. 1.5e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 AAAAATCTCTAGGCTTCAGT 24
Db 43155 AAAAATCTCTAGGCTTCATT 43132
RESULT 23
AAK84730/c
ID AAK84730 standard; DNA; 48045 BP.
XX
AC AAK84730;
XX
DT 07-NOV-2001 (first entry)
XX
DE Human immune/hematopoietic antigen genomic sequence SEQ ID NO:39542.
XX
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX
OS Homo sapiens.
XX
FN WO200157182-A2.
XX
PD 09-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001354.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216800P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.

RESULT 24	PR	08-SEP-2000;	2000US-0232081P.
AAK85984/c	PR	12-SEP-2000;	2000US-0231968P.
ID AAK85984 standard; DNA; 48045 BP.	PR	14-SEP-2000;	2000US-0232397P.
XX	PR	14-SEP-2000;	2000US-0232398P.
AC AAK85984;	PR	14-SEP-2000;	2000US-0232399P.
XX	PR	14-SEP-2000;	2000US-0232400P.
XX	PR	14-SEP-2000;	2000US-0232401P.
DT 07-NOV-2001 (first entry)	PR	14-SEP-2000;	2000US-0233063P.
XX	PR	14-SEP-2000;	2000US-0233064P.
XX	PR	14-SEP-2000;	2000US-0233065P.
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:40796.	PR	21-SEP-2000;	2000US-0234223P.
XX	PR	21-SEP-2000;	2000US-0234274P.
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;	PR	25-SEP-2000;	2000US-0234997P.
KW cytostatic; gene therapy; vaccine; metastasis; ds.	PR	25-SEP-2000;	2000US-0234998P.
XX	PR	26-SEP-2000;	2000US-0235484P.
OS Homo sapiens.	PR	27-SEP-2000;	2000US-0235834P.
XX	PR	27-SEP-2000;	2000US-0235836P.
PN WO200157182-A2.	PR	29-SEP-2000;	2000US-0236327P.
XX	PR	29-SEP-2000;	2000US-0236367P.
PD	PR	29-SEP-2000;	2000US-0236368P.
XX	PR	29-SEP-2000;	2000US-0236369P.
PF	PR	29-SEP-2000;	2000US-0236370P.
XX	PR	02-OCT-2000;	2000US-0236802P.
XX	PR	02-OCT-2000;	2000US-0237037P.
PR	PR	02-OCT-2000;	2000US-0237038P.
PR	PR	02-OCT-2000;	2000US-0237039P.
PR	PR	02-OCT-2000;	2000US-0237040P.
PR	PR	13-OCT-2000;	2000US-0239935P.
PR	PR	13-OCT-2000;	2000US-0239937P.
PR	PR	20-OCT-2000;	2000US-0240960P.
PR	PR	20-OCT-2000;	2000US-0241221P.
PR	PR	20-OCT-2000;	2000US-0241785P.
PR	PR	20-OCT-2000;	2000US-0241786P.
PR	PR	20-OCT-2000;	2000US-0241787P.
PR	PR	20-OCT-2000;	2000US-0241808P.
PR	PR	20-OCT-2000;	2000US-0241809P.
PR	PR	20-OCT-2000;	2000US-0241826P.
PR	PR	01-NOV-2000;	2000US-0244617P.
PR	PR	08-NOV-2000;	2000US-0246474P.
PR	PR	08-NOV-2000;	2000US-0246475P.
PR	PR	08-NOV-2000;	2000US-0246476P.
PR	PR	08-NOV-2000;	2000US-0246477P.
PR	PR	08-NOV-2000;	2000US-0246478P.
PR	PR	08-NOV-2000;	2000US-0246523P.
PR	PR	08-NOV-2000;	2000US-0246524P.
PR	PR	08-NOV-2000;	2000US-0246525P.
PR	PR	08-NOV-2000;	2000US-0246526P.
PR	PR	08-NOV-2000;	2000US-0246527P.
PR	PR	08-NOV-2000;	2000US-0246528P.
PR	PR	08-NOV-2000;	2000US-0246532P.
PR	PR	08-NOV-2000;	2000US-0246609P.
PR	PR	08-NOV-2000;	2000US-0246610P.
PR	PR	08-NOV-2000;	2000US-0246611P.
PR	PR	08-NOV-2000;	2000US-0246613P.
PR	PR	17-NOV-2000;	2000US-0249207P.
PR	PR	17-NOV-2000;	2000US-0249208P.
PR	PR	17-NOV-2000;	2000US-0249209P.
PR	PR	17-NOV-2000;	2000US-0249210P.
PR	PR	17-NOV-2000;	2000US-0249211P.
PR	PR	17-NOV-2000;	2000US-0249212P.
PR	PR	17-NOV-2000;	2000US-0249213P.
PR	PR	17-NOV-2000;	2000US-0249214P.
PR	PR	17-NOV-2000;	2000US-0249215P.
PR	PR	17-NOV-2000;	2000US-0249216P.
PR	PR	17-NOV-2000;	2000US-0249217P.
PR	PR	17-NOV-2000;	2000US-0249218P.
PR	PR	17-NOV-2000;	2000US-0249244P.
PR	PR	17-NOV-2000;	2000US-0249245P.
PR	PR	17-NOV-2000;	2000US-0249264P.
PR	PR	17-NOV-2000;	2000US-0249265P.
PR	PR	17-NOV-2000;	2000US-0249297P.
PR	PR	17-NOV-2000;	2000US-0249299P.
PR	PR	17-NOV-2000;	2000US-0249300P.

```
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX
XX Rosen CA, Barash SC, Ruben SM;
PI WPI; 2001-483426/52.
XX
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and metastasis.
XX
XX Disclosure; SEQ ID NO 40796; 3071pp + Sequence Listing; English.
XX
XX AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC amino acid sequences given in AAK82170 to AAK51921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patient's own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting the
CC nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
CC represent sequences used in the exemplification of the present invention
XX
XX Sequence 48045 BP; 12889 A; 9812 C; 10461 G; 14883 T; 0 U; 0 Other;
SQ
Query Match 76.8%; Score 19.2; DB 4; Length 48045;
Best Local Similarity 87.5%; Pred. No. 1.5e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 AAAAACTCTGTAGGCTTTTCAGT 24
Db 43163 AAAAAATTCCTGTAGGTTTTCATT 43140
RESULT 25
ADD33112
ID ADD33112 standard; DNA; 466 BP.
XX
XX ADD33112;
XX
XX 15-JAN-2004 (first entry)
XX
XX Human mitochondrial DNA sequence SEQ ID NO:883.
XX
XX ds; human; array; mitochondrial; hybridisation; energy-metabolism;
KW mitochondrial disease; oxidative phosphorylation dysfunction;
KW oxidative stress; apoptosis; aging.
XX
XX Homo sapiens.
XX
XX WO2003020220-A2.
XX
XX 13-MAR-2003.
PD
```

```
XX 30-AUG-2002; 2002WO-US027886.
XX
XX 30-AUG-2001; 2001US-0316323P.
PR 31-AUG-2001; 2001CA-02356540.
XX
XX (UYEM-) UNIV EMORY.
XX
XX Wallace DC, Levy S, Kerstann K, Procaccio V;
PI WPI; 2003-300821/29.
XX
XX Array containing probes for genes involved in mitochondrial biology,
PT useful for determining mitochondrial biology gene expression profiles for
PT use in diagnosing pathologies and identifying biochemical pathways.
XX
XX Claim 2; SEQ ID NO 883; 201pp; English.
XX
XX The invention relates to a novel array comprising at least two isolated
CC nucleotide molecules, each molecule having a sequence capable of uniquely
CC hybridising to a nucleic acid molecule which is an expression product of
CC a gene involved in mitochondrial biology. The array comprises two or more
CC isolated nucleic acid molecules or spots, each molecule having a sequence
CC chosen from sequence of 994 human probes and 2046 mouse probes. An array
CC of the invention is useful for determining an expression profile of a
CC mouse or human sample containing nucleic acid, by contacting the array
CC with the sample under conditions allowing selective hybridisation, and
CC measuring hybridisation of nucleic acid in the sample to the array to
CC produce an expression profile. The array is also useful for determining
CC an expression profile of a first labelled sample containing nucleic acid
CC relative to a second, differently labelled sample containing nucleic
CC acid. The second sample is a reference or a standard. An array is useful
CC for determining an expression profile diagnostic of an energy-metabolism-
CC related physiological condition. An array of the invention is useful for
CC determining mitochondrial biology gene expression profiles of organisms,
CC such as human, mice and closely related species, tissue and organs of
CC such organisms, which are useful for determining expression profiles
CC diagnostic of energy metabolism-related physiological conditions,
CC diagnosing such physiological conditions, identifying biochemical
CC pathways, genes, and mutations involved in such physiological conditions,
CC identifying therapeutic agents useful for preventing and/or treating such
CC physiological conditions, evaluating and/or monitoring the efficacy of
CC such therapies, and creating and identifying animal models of human
CC energy metabolism-related physiological conditions. An array is also
CC useful for defining expression signatures or profiles for mitochondrial
CC diseases, as well as distinguishing clinical disorders that result from
CC oxidative phosphorylation (OXPHOS) dysfunction, oxidative stress,
CC apoptosis and aging. An array of the invention contains probes of genes
CC not previously recognised to participate in mitochondrial biology. The
CC sequences shown in ADD32231-ADD33223 represent human mitochondrial DNA
CC clones used to make the probes of the invention. Some sequences are not
CC present, these are SEQ ID NO's 295, 1174, 1213, 1700, 1728, 1730, 1905,
XX 1906, 2408 and 2643.
XX
XX Sequence 466 BP; 86 A; 122 C; 124 G; 130 T; 0 U; 4 Other;
SQ
Query Match 74.4%; Score 18.6; DB 10; Length 466;
Best Local Similarity 84.0%; Pred. No. 1.4e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 AAAAACTCTGTAGGCTTTTCAGT 25
Db 328 AAATACATCCTGTAGGCTTTTCGTG 352
RESULT 26
AAD48290
ID AAD48290 standard; DNA; 39982 BP.
XX
XX AAD48290;
XX
XX 24-FEB-2003 (first entry)
XX
```


FT variation replace(15840, -)
FT /*tag= ay
FT /standard name= "Single nucleotide polymorphism"
FT replace(17726,A)
FT /*tag= az
FT /standard name= "Single nucleotide polymorphism"
FT replace(17736,G)
FT /*tag= ba
FT /standard name= "Single nucleotide polymorphism"
FT replace(18850,T)
FT /*tag= bb
FT /standard name= "Single nucleotide polymorphism"
FT replace(19419, -)
FT /*tag= bc
FT /standard name= "Single nucleotide polymorphism"
FT replace(19429,G)
FT /*tag= bd
FT /standard name= "Single nucleotide polymorphism"
FT replace(19431,G)
FT /*tag= be
FT /standard name= "Single nucleotide polymorphism"
FT 21594..21735
FT /*tag= bf
FT 21736..25190
FT /*tag= bg
FT variation replace(22516,T)
FT /*tag= bh
FT /standard name= "Single nucleotide polymorphism"
FT replace(22543,T)
FT /*tag= bi
FT /standard name= "Single nucleotide polymorphism"
FT replace(23568,A)
FT /*tag= bj
FT /standard name= "Single nucleotide polymorphism"
FT replace(23779,A)
FT /*tag= bk
FT /standard name= "Single nucleotide polymorphism"
FT replace(23808,A)
FT /*tag= bl
FT /standard name= "Single nucleotide polymorphism"
FT replace(23979,C)
FT /*tag= bm
FT /standard name= "Single nucleotide polymorphism"
FT replace(24319,C)
FT /*tag= bn
FT /standard name= "Single nucleotide polymorphism"
FT 25191..25302
FT /*tag= bo

Query Match 74.4%; Score 18.6; DB 8; Length 39982;
Best Local Similarity 84.0%; Pred. No. 2.7e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAACTCCTGTAGGCTTTTCAGTG 25
|||||
Db 14081 AAGAAATCCTGTGGCTTTTCGTG 14105

RESULT 27
ID ADL13771
ADL13771 standard; DNA; 201766 BP.

XX AC ADL13771;

XX 06-MAY-2004 (first entry)

XX Osteoarthritis-associated polymorphic nucleotide #303.

XX ds; gene; osteopathic; antiinflammatory; antiarthritic; gene therapy;
KW joint space narrowing; osteophyte development; joint pain;
KW osteoarthritis; SNP; single nucleotide polymorphism.
XX
OS Homo sapiens.

XX WO2003054166-A2.
XX 03-JUL-2003.
XX 19-DEC-2002; 2002WO-US041225.
XX 20-DEC-2001; 2001US-0342603P.
XX (INCY-) INCYTE GENOMICS INC.
XX Jones KA, Schafer A;
XX WPI; 2003-559141/52.
XX
XX Determining susceptibility of an individual to joint space narrowing,
XX osteophyte development and/or joint pain comprises identifying whether
XX the individual has at least one polymorphism in a polynucleotide encoding
XX a protein.
XX
XX Disclosure; SEQ ID NO 303; 297pp; English.
XX
XX The invention relates to a method of determining susceptibility of an
XX individual to joint space narrowing and/or osteophyte development and/or
XX joint pain comprising identifying whether the individual has at least one
XX polymorphism in a polynucleotide encoding at least one of the protein
XX listed in the specification. The methods, composition and agent are
XX useful for modulating the susceptibility of an individual to joint space
XX narrowing and/or osteophyte development and/or joint pain that is
XX associated with a disease, preferably osteoarthritis. The cell line and
XX the non-human animal are useful for screening for an agent for diagnosing
XX an individual having susceptibility to joint space narrowing and/or
XX osteophyte development and/or joint pain. This sequence corresponds to
XX the polynucleotide encoding a protein listed in the specification. (Note:
XX The sequence data for this patent did not form part of the printed
XX specification but was obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences).
XX
XX Sequence 201766 BP; 57593 A; 35575 C; 36821 G; 54688 T; 0 U; 17089 Other;

Query Match 74.4%; Score 18.6; DB 10; Length 201766;
Best Local Similarity 84.0%; Pred. No. 3.5e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAACTCCTGTAGGCTTTTCAGTG 25
|||||
Db 24470 AAAAACTACTATAGGCTTCAGTG 24494

RESULT 28
AAC36832/c
ID AAC36832 standard; DNA; 1175 BP.

XX AC AAC36832;

XX 17-OCT-2000 (first entry)

XX Arabidopsis thaliana DNA fragment SEQ ID NO: 15227.

XX Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX Arabidopsis thaliana.

XX EP1033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-00301439.

XX 25-FEB-1999; 99US-0121825P.

XX 05-MAR-1999; 99US-0123180P.

PR

PR 09-MAR-1999; 99US-0123548B.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132407P.
PR 05-MAY-1999; 99US-0132484P.
PR 06-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132487P.
PR 11-MAY-1999; 99US-0132863P.
PR 14-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 18-MAY-1999; 99US-0134370P.
PR 19-MAY-1999; 99US-0134768P.
PR 20-MAY-1999; 99US-0134941P.
PR 21-MAY-1999; 99US-0135124P.
PR 24-MAY-1999; 99US-0135353P.
PR 25-MAY-1999; 99US-0135629P.
PR 27-MAY-1999; 99US-0136021P.
PR 28-MAY-1999; 99US-0136392P.
PR 01-JUN-1999; 99US-0136782P.
PR 03-JUN-1999; 99US-0137222P.
PR 04-JUN-1999; 99US-0137528P.
PR 07-JUN-1999; 99US-0137502P.
PR 08-JUN-1999; 99US-0137724P.
PR 10-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142377P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-014902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.

```
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 22-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161933P.
PR 29-OCT-1999; 99US-0162142P.
PR 29-OCT-1999; 99US-0162143P.
PR 29-OCT-1999; 99US-0162144P.

Query Match 73.6%; Score 18.4; DB 3; Length 1175;
Best Local Similarity 95.0%; Pred. No. 2e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 AACTCCTGTAGGCTTTCAGT 24
   |||||
Db 367 AACTCCTGTAGGCTTTCAGT 348

RESULT 29
AAC42714/c
ID AAC42714 standard; DNA; 2286 BP.
XX
AC AAC42714;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 36582.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
```

PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 21-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.

PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 22-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 73.6%; Score 18.4; DB 3; Length 2286;
Best Local Similarity 95.0%; Pred. NO. 2.2e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 AACTCCTGTAGCGTTTCAGT 24
|||||
Db 1550 AACTCCTGTAGCGTTTCAGT 1531

RESULT 30
ADN74670/c
ID ADN74670 standard; cDNA; 2286 BP.
XX AC ADN74670;
XX 15-JUL-2004 (first entry)
XX DE Thale cress cDNA repressed in E2Fa/Dpa expressing plants SeqID 2565.
XX KW gene; ss; plant; transgenic; E2Fa/Dpa transcription factor;
KW growth regulator; animal feed product; thale cress;
KW cell wall biosynthesis; nitrogen metabolism; carbon metabolism.
XX OS Arabidopsis thaliana.
XX PN WO2004035798-A2.
XX PD 29-APR-2004.
XX PF 20-OCT-2003; 2003WO-EP011658.
XX PR 18-OCT-2002; 2002EP-00079408.
XX PA (CROP-) CROPDESIGN NV.
XX PI Inze D, De Veylder L, Vlieghe K;
XX WPI; 2004-348466/32.
XX P-PSDB; ADN74671.
XX Altering plant characteristics, useful for producing plants for enzyme or
XX pharmaceutical production comprising modifying in a plant, expression of
XX one or more nucleic acids and/or modifying level or activity of one or
XX more proteins.
XX Claim 1; SEQ ID NO 2565; 134pp; English.

CC This invention relates to a novel method for altering one or more plant
CC characteristics. Specifically, it refers to identifying genes that are up
CC - or down-regulated in transgenic plants overexpressing the heterodimeric
CC E2Fa/DPA transcription factor of Arabidopsis and using these sequences to
CC alter plant characteristics accordingly. The present invention describes
CC generating transgenic plants for the production of growth regulators,
CC enzymes, therapeutics, pharmaceuticals and animal feed products, where
CC the altered plant characteristics are selected from increased yield or
CC biomass, enhanced survival capacity, stress tolerance, plant architecture
CC or physiology, altered endoreduplication, biochemistry, signal
CC transduction, storage lipid mobilization and/or altered photosynthesis,
CC each relative to the corresponding wild type plants. Accordingly, these
CC sequences can also be useful as positive or negative selectable markers
CC during transformation of cells or tissues. The identified genes play a
CC role in a variety of biological processes such as DNA replication, cell
CC wall biosynthesis, nitrogen and/or carbon metabolism or they function as
CC transcription factors. This polynucleotide sequence is thale cress cDNA
CC repressed 1.3 fold or more in plants overexpressing the E2Fa/DPA
CC transcription factor, given in an exemplification of the invention.

SQ Sequence 2286 BP; 535 A; 544 C; 462 G; 745 T; 0 U; 0 Other;

Query Match 73.6%; Score 18.4; DB 12; Length 2286;
Best Local Similarity 95.0%; Pred. No. 2.2e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 AACTCTGTAGGCTTTCAGT 24
|||||
DB 1550 AACTCTGTAGGCTTTCAGT 1531

RESULT 31
ABN22936
ID ABN22936 standard; cDNA; 463 BP.

XX AC ABN22936;

XX DT 24-JUN-2002 (first entry)

XX DE Human ORFX polynucleotide sequence SEQ ID NO:14349.

XX KW Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;
KW hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;
KW degenerative disorder; osteoarthritis; neurodegenerative disorder;
KW cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;
KW hypertension; hypothyroidism; cholesterol ester storage disease;
KW immune deficiency; immune disorder; infectious disease;
KW autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;
KW myasthenia gravis; gene; ss.

XX OS Homo sapiens.

XX XX WO2001192523-A2.

XX PN 06-DEC-2001.

XX PD 29-MAY-2001; 2001WO-US010836.

XX PF 30-MAY-2000; 2000US-0206132P.

XX PR 29-AUG-2000; 2000US-0228716P.

XX XX (CURA-) CURAGEN CORP.

XX PA Shimketa RA, Leach MD;

XX PI WPI; 2002-106308/14.

XX XX P-PSDB; ABP07184.

XX XX Novel human polypeptides and polynucleotides useful for diagnosing,
PT preventing and treating cardiovascular disease, neurodegenerative,
PT hyperproliferative disorders and autoimmune disorders.

XX PS Disclosure; SEQ ID NO 14349; 1037pp; English.

XX CC The present invention describes substantially purified human proteins
CC (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1
CC in the specification). ABN15762 to ABN27252 encode the human ORFX
CC proteins given in ABP00010 to ABP11500. ORFX proteins are useful for
CC treating or preventing a pathology associated with an ORFX-associated
CC disorder in humans, and in the manufacture of a medicament for treating a
CC syndrome associated with ORFX-associated disorder. ORFX polynucleotide
CC sequences can be used in gene therapy. ORFX sequences can be used in the
CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver,
CC psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,
CC osteoarthritis, neurodegenerative disorders, disorders related to organ
CC transplantation, cardiovascular diseases, diabetes mellitus, systemic
CC lupus erythematosus, hypertension, hypothyroidism, cholesterol ester
CC storage disease, various immune deficiencies and disorders, infectious
CC disease, autoimmune disorders such as multiple sclerosis, rheumatoid
CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host
CC disease and autoimmune inflammatory eye disease. ORFX proteins are also
CC useful for treating burns, incisions, ulcers, for treating osteoporosis,
CC bone degenerative disorders, or periodontal disease, and for gut
CC protection or regeneration and treatment of lung or liver fibrosis,
CC repulsion or regeneration in various tissues and conditions resulting from
CC systemic cytokine damage. N.B. The sequence data for this patent did not
CC form part of the printed specification, but was obtained in electronic
CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX CC
SQ Sequence 463 BP; 136 A; 100 C; 71 G; 156 T; 0 U; 0 Other;

Query Match 72.8%; Score 18.2; DB 6; Length 463;
Best Local Similarity 87.0%; Pred. No. 2.1e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAATCTCTAGGCTTTCAGT 24
|||||
DB 198 AAAAATCTCTAGGCTTTCAGT 220

RESULT 32

ADH10017_1

Continuation (2 of 4) of ADH10017 from base 100001 (Human chromosome 2p21-22 fragment
WP Sequence split into 4 fragments LOCUS ADH10017 Accession Adh10017

WP	Fragment Name	Begin	End
WP	ADH10017_0	1	110000
WP	ADH10017_1	100001	210000
WP	ADH10017_2	200001	310000
WP	ADH10017_3	300001	365186

Query Match 72.8%; Score 18.2; DB 10; Length 110000;
Best Local Similarity 87.0%; Pred. No. 4.9e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAATCTCTAGGCTTTCAGT 24
|||||
DB 24978 AAAAATCTCTAGGCTTTCAGT 25000

RESULT 33

ADQ97266_0

WP Sequence split into 4 fragments LOCUS ADQ97266 Accession Adq97266

WP	Fragment Name	Begin	End
WP	ADQ97266_0	1	110000
WP	ADQ97266_1	100001	210000
WP	ADQ97266_2	200001	310000
WP	ADQ97266_3	300001	355211

ID ADQ97266 standard; DNA; 355211 BP.

XX AC ADQ97266;

XX DT 07-OCT-2004 (first entry)

XX DE Human cancer associated sequence HD08-023, SEQ ID 242.

XX KW Cytostatic; Gene Therapy; cancer; leukemia; lymphoma; Human; ds.

```
XX OS Homo sapiens.
XX PN WO2004060304-A2.
XX PD 22-JUL-2004.
XX PF 22-DEC-2003; 2003WO-US041389.
XX PR 27-DEC-2002; 2002US-00330773.
XX PA (SAGR-) SAGRES DISCOVERY INC.
XX PI Morris DW, Malandro MS;
XX DR WPI; 2004-543781/52.
XX PT New isolated cancer associated nucleic acids comprising at least 10
XX PT contiguous nucleotides, useful for diagnosing, preventing and/or treating
XX PT cancers such as leukemia and lymphoma.
XX PS Claim 1; SEQ ID NO 242; 199pp; English.
XX CC The present invention relates to cancer associated sequences (ADQ97025-
XX CC ADQ98004). The sequences are useful for the diagnosis, prevention and/or
XX CC treatment of cancer, such as leukemia and lymphoma. Note: The sequence
XX CC data for this patent did not form part of the printed specification, but
XX CC was obtained in electronic formate directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 355211 BP; 96742A; 68188C; 72417G; 106131T; 0U; 117330Other;

Query Match 72.8%; Score 18.2; DB 12; Length 110000;
Best Local Similarity 87.0%; Pred. No. 4.9e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAACTCCTGTAGGCTTTTCAGT 24
Db 105281 AAAAAACACCTGTATGTTTTCAGT 105303

RESULT 34
ADQ97266_1
Continuation (2 of 4) of ADQ97266 from base 100001 (Human cancer associated sequence HDQ
WP Sequence split into 4 fragments LOCUS ADQ97266 Accession Adq97266
WP Fragment Name Begin End
WP ADQ97266_0 1 110000
WP ADQ97266_1 100001 210000
WP ADQ97266_2 200001 310000
WP ADQ97266_3 300001 355211

Query Match 72.8%; Score 18.2; DB 12; Length 110000;
Best Local Similarity 87.0%; Pred. No. 4.9e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAACTCCTGTAGGCTTTTCAGT 24
Db 5281 AAAAAACACCTGTATGTTTTCAGT 5303

RESULT 35
ADQ97328_2/c
Continuation (3 of 4) of ADQ97328 from base 200001 (Mouse cancer associated sequence MDQ
WP Sequence split into 4 fragments LOCUS ADQ97328 Accession Adq97328
WP Fragment Name Begin End
WP ADQ97328_0 1 110000
WP ADQ97328_1 100001 210000
WP ADQ97328_2 200001 310000
WP ADQ97328_3 300001 358847

Query Match 72.8%; Score 18.2; DB 12; Length 110000;
Best Local Similarity 87.0%; Pred. No. 4.9e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 2 AAAAACTCCTGTAGGCTTTTCAGT 24
Db 97821 AAAAAATCCCTGTAGTCTTTTCAGT 97799

RESULT 36
ABD33491/c
ID ABD33491 standard; DNA; 112241 BP.
XX AC ABD33491;
XX DT 18-NOV-2004 (first entry)
XX DE Human cancer-associated (CA) gene HD07-095.
XX KW Human; cancer-associated protein; CAP; cancer-associated gene; CA; gene;
XX KW ds; cancer; cytostatic.
XX OS Homo sapiens.
XX PN WO2004058146-A2.
XX PD 15-JUL-2004.
XX PF 15-DEC-2003; 2003WO-US040081.
XX PR 17-DEC-2002; 2002US-00322281.
XX PA (SAGR-) SAGRES DISCOVERY INC.
XX PI Morris DW, Malandro MS;
XX DR WPI; 2004-499109/47.
XX PT Novel human cancer associated protein encoded within open reading frame
XX PT of cancer associated gene, useful as targets for diagnosing cancer.
XX PS Claim 16; SEQ ID NO 656; 182pp; English.
XX CC The invention relates to cancer-associated proteins (CAP) and the cancer-
XX CC associated (CA) nucleic acids encoding them. The invention also relates
XX CC to a method for treating cancers involving administering to a patient an
XX CC inhibitor of CAP, and a method of screening for anticancer activity in a
XX CC potential drug involving providing a cell that expresses a CA gene
XX CC contacting a tissue sample derived from a cancer cell with an anticancer
XX CC drug candidate and monitoring the effect of the anticancer drug candidate
XX CC on expression of the CA gene. The CAP proteins are useful for detecting
XX CC cancer associated with expression of a CAP protein in a test cell sample
XX CC and for screening for a bioactive agent capable of modulating the
XX CC activity of a CAP protein. The CA nucleic acids are useful for diagnosing
XX CC cancer, involving determining the expression of a CA nucleic acid in a
XX CC tissue. This sequence represents a human CA gene of the invention. Note:
XX CC The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 112241 BP; 28197 A; 21716 C; 21810 G; 31271 T; 0 U; 9247 Other;

Query Match 72.8%; Score 18.2; DB 13; Length 112241;
Best Local Similarity 87.0%; Pred. No. 5e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAACTCCTGTAGGCTTTTCAGT 24
Db 84160 AAAAAATCCCTGTGTTGCTTTTCAGT 84138

RESULT 37
ADL13676
ID ADL13676 standard; DNA; 192639 BP.
XX AC ADL13676;
```

XX 06-MAY-2004 (first entry)
DT Osteoarthritis-associated polymorphic nucleotide #208.
XX ds; gene; osteopathic; antiinflammatory; antiarthritic; gene therapy;
KW joint space narrowing; osteophyte development; joint pain;
KW osteoarthritis; SNP; single nucleotide polymorphism.
XX Homo sapiens.
OS WO2003054166-A2.
XX 03-JUL-2003.
XX 19-DEC-2002; 2002WO-US041225.
PF 20-DEC-2001; 2001US-0342603P.
XX (INCY-) INCYTE GENOMICS INC.
PA Jones KA, Schafer A;
XX WPI; 2003-559141/52.
XX Determining susceptibility of an individual to joint space narrowing.
PT osteophyte development and/or joint pain comprises identifying whether
PT the individual has at least one polymorphism in a polynucleotide encoding
PT a protein.
XX Disclosure; SEQ ID NO 208; 297pp; English.
PS The invention relates to a method of determining susceptibility of an
CC individual to joint space narrowing and/or osteophyte development and/or
CC joint pain comprising identifying whether the individual has at least one
CC polymorphism in a polynucleotide encoding at least one of the protein
CC listed in the specification. The methods, composition and agent are
CC useful for modulating the susceptibility of an individual to joint space
CC narrowing and/or osteophyte development and/or joint pain that is
CC associated with a disease, preferably osteoarthritis. The cell line and
CC the non-human animal are useful for screening for an agent for diagnosing
CC an individual having susceptibility to joint space narrowing and/or
CC osteophyte development and/or joint pain. This sequence corresponds to
CC the polynucleotide encoding a protein listed in the specification. (Note:
CC The sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences).
XX SQ Sequence 192639 BP; 56569 A; 40174 C; 37841 G; 57918 T; 0 U; 137 Other;
Query Match 72.8%; Score 18.2; DB 10; Length 192639;
Best Local Similarity 87.0%; Pred. No. 5.3e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 3 AAAAATCTCTAGGCTTCAGTG 25
Db 150038 AAAAATCTCTAGGCTTCAGTG 150060
RESULT 38
AAN82201/c
ID AAN82201 standard; DNA; 1794 BP.
XX AAN82201;
AC 25-MAR-2003 (revised)
DT 10-MAR-2003 (revised)
DT 05-DEC-1990 (first entry)
XX Beta-amylase from plant seed.
DE Beta-amylase; plant seed; leader sequence; soybean; ss.
KW Beta-amylase; plant seed; leader sequence; soybean; ss.
XX

OS Glycine max.
XX Key Location/Qualifiers
FH 108..1595
FT CDS /*tag=a
FT /*product="beta-amylase"
XX JP63091084-A.
XX 21-APR-1988.
XX 06-OCT-1986; 86JP-00236341.
XX 06-OCT-1986; 86JP-00236341.
XX (NORQ) NORINSHO KK.
XX WPI; 1988-150872/22.
DR P-PSDB; AAP82687.
XX Beta-amylase messenger RNA and novel recombinant - obtd. from plant seed
PT by fractionation by saccharose density slope.
XX Disclosure; Page ?; 6pp; Japanese.
XX The messenger RNA corresp. to beta-amylase, obtd. from plant seed and
CC free from leader sequence is obtd. in a fraction heavier than 28S by
CC fractionation by saccharose density slope centrifugal sepn. method. Beta-
CC amylase of soybean has heat resistance higher than beta-amylase from
CC microorganisms. (Updated on 10-MAR-2003 to add missing OS field.)
CC (Updated on 25-MAR-2003 to correct PA field.)
XX SQ Sequence 1794 BP; 558 A; 372 C; 385 G; 479 T; 0 U; 0 Other;
Query Match 71.2%; Score 17.8; DB 1; Length 1794;
Best Local Similarity 90.5%; Pred. No. 4e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 AAAAATCTCTAGGCTTC 21
Db 337 AACAAATCTCTAGGCTTC 317
RESULT 39
ABQ69245_03/c
Continuation (4 of 31) of ABQ69245 from base 300001 (Listeria innocua DNA sequence #684
WP Sequence split into 31 fragments LOCUS ABQ69245 Accession Abq69245
WP Fragment Name Begin End
WP ABQ69245_00 1 110000
WP ABQ69245_01 100001 210000
WP ABQ69245_02 200001 310000
WP ABQ69245_03 300001 410000
WP ABQ69245_04 400001 510000
WP ABQ69245_05 500001 610000
WP ABQ69245_06 600001 710000
WP ABQ69245_07 700001 810000
WP ABQ69245_08 800001 910000
WP ABQ69245_09 900001 1010000
WP ABQ69245_10 1000001 1110000
WP ABQ69245_11 1100001 1210000
WP ABQ69245_12 1200001 1310000
WP ABQ69245_13 1300001 1410000
WP ABQ69245_14 1400001 1510000
WP ABQ69245_15 1500001 1610000
WP ABQ69245_16 1600001 1710000
WP ABQ69245_17 1700001 1810000
WP ABQ69245_18 1800001 1910000
WP ABQ69245_19 1900001 2010000
WP ABQ69245_20 2000001 2110000
WP ABQ69245_21 2100001 2210000
WP ABQ69245_22 2200001 2310000
WP ABQ69245_23 2300001 2410000
WP ABQ69245_24 2400001 2510000

```
WP ABQ69245_25 2500001 2610000
WP ABQ69245_26 2600001 2710000
WP ABQ69245_27 2700001 2810000
WP ABQ69245_28 2800001 2910000
WP ABQ69245_29 2900001 3010000
WP ABQ69245_30 3000001 3110000

Query Match 71.2%; Score 17.8; DB 6; Length 110000;
Best Local Similarity 90.5%; Pred. No. 7.6e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AAAAATCCTGTAGGCTTTCA 22
Db 13500 AAAAATCCTGTAGGCTTTCA 13480

RESULT 40
ABQ67197_02/c
Continuation (3 of 12) of ABQ67197 from base 200001 (Listeria innocua contig DNA sequen
WP Sequence split into 12 fragments LOCUS ABQ67197 Accession Abq67197
WP Fragment Name Begin End
WP ABQ67197_00 1 110000
WP ABQ67197_01 100001 210000
WP ABQ67197_02 200001 310000
WP ABQ67197_03 300001 410000
WP ABQ67197_04 400001 510000
WP ABQ67197_05 500001 610000
WP ABQ67197_06 600001 710000
WP ABQ67197_07 700001 810000
WP ABQ67197_08 800001 910000
WP ABQ67197_09 900001 1010000
WP ABQ67197_10 1000001 1110000
WP ABQ67197_11 1100001 1163020

Query Match 71.2%; Score 17.8; DB 6; Length 110000;
Best Local Similarity 90.5%; Pred. No. 7.6e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AAAAATCCTGTAGGCTTTCA 22
Db 23235 AAAAATCCTGTAGGCTTTCA 23215

RESULT 41
AAL47655/c
ID AAL47655 standard; DNA; 43 BP.
XX
AC AAL47655;
XX
DT 18-SEP-2002 (first entry)
XX
DE Specific sequence isolation/determination method anchor primer #7.
XX
KW Gene specific sequence isolation; sequence determination; disease;
KW diagnosis; infection; forensics; paternity testing; PCR; primer; anchor;
XX ss.
XX Synthetic.
XX Key Location/Qualifiers
FT modified_base 1 /*tag= a
FT /*mod_base= OTHER
FT /*note= "5' biotinylated"
XX
PN WO200183696-A2.
XX
PD 08-NOV-2001.
XX
PF 27-APR-2001; 2001WO-US013807.
XX
PR 28-APR-2000; 2000US-00560845.
XX
```

```
PA (DIGI-) DIGITAL GENE TECHNOLOGIES INC.
XX
PI Muller R, Riddle GH, Glass JR;
XX
DR WPI; 2002-114150/15.
XX
PT Amplifying gene-specific(GS) polynucleotide by producing population of
PT DNA molecules having adapter molecule and anchor primer that has
PT captureable group, synthesizing and amplifying single-stranded GS
PT polynucleotide.
XX
PS Claim 68; Page 102; 115pp; English.
XX
CC The present invention relates to a method of isolating, amplifying and
CC sequencing gene specific sequences. This involves (a) synthesising double
CC -stranded (ds) DNAs containing the gene specific polynucleotide which has
CC an anchor primer with a captureable group at the 3' end; (b) ligating a
CC ds adapter molecule to each dsDNA; (c) synthesising a single-stranded
CC (ss) gene specific polynucleotide using a gene specific oligonucleotide
CC primer having a second captureable group; and (d) amplifying the single
CC stranded gene specific polynucleotide after isolation. The method is used
CC for determining without cloning an extended sequence of a gene specific
CC polynucleotide that has a partial gene sequence known by carrying out the
CC method and then sequencing the amplified isolated gene specific
CC polynucleotide without cloning. The method has several diagnostic and
CC research applications, including genetic and infectious disease
CC diagnosis, forensic and paternity testing, toxicology testing, individual
CC identification, genetic research, occupational hazard screening and
CC pharmaceutical development. The method can be used to detect mutations in
CC target genes which may lead to disease conditions such as Alzheimer's
CC disease, cancer and Duchenne muscular dystrophy, thus facilitating the
CC diagnosis of genetic diseases as well as the identification of carriers
CC of recessive genetic disorders. The present sequence is an
CC oligonucleotide useful in the method of the invention
XX
SQ Sequence 43 BP; 10 A; 4 C; 7 G; 22 T; 0 U; 0 Other;

Query Match 70.4%; Score 17.6; DB 6; Length 43;
Best Local Similarity 83.3%; Pred. No. 2.8e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAATCCTGTAGGCTTTCA 24
Db 31 AAAAATCCTGTAGGCTTTCA 8

RESULT 42
AAL47656/c
ID AAL47656 standard; DNA; 43 BP.
XX
AC AAL47656;
XX
DT 18-SEP-2002 (first entry)
XX
DE Specific sequence isolation/determination method anchor primer #8.
XX
KW Gene specific sequence isolation; sequence determination; disease;
KW diagnosis; infection; forensics; paternity testing; PCR; primer; anchor;
XX ss.
XX Synthetic.
XX Key Location/Qualifiers
FT modified_base 1 /*tag= a
FT /*mod_base= OTHER
FT /*note= "5' aminated"
XX
PN WO200183696-A2.
XX
PD 08-NOV-2001.
XX
PF 27-APR-2001; 2001WO-US013807.
XX
```


XX 28-APR-2000; 2000US-00560845.
PR (DIGI-) DIGITAL GENE TECHNOLOGIES INC.
XX Muller R, Riddle GH, Glass JR;
PI WPI; 2002-114150/15.
XX Amplifying gene-specific(GS) polynucleotide by producing population of
PT DNA molecules having adapter molecule and anchor primer that has
PT captureable group, synthesizing and amplifying single-stranded GS
PT polynucleotide.
XX Claim 68; Page 103; 115pp; English.
XX The present invention relates to a method of isolating, amplifying and
CC sequencing gene specific sequences. This involves (a) synthesizing double
CC -stranded (ds) DNAs containing the gene specific polynucleotide which has
CC an anchor primer with a captureable group at the 3' end; (b) ligating a
CC ds adapter molecule to each dsDNA; (c) synthesizing a single-stranded
CC (ss) gene specific polynucleotide using a gene specific oligonucleotide
CC primer having a second captureable group; and (d) amplifying the single
CC stranded gene specific polynucleotide after isolation. The method is used
CC for determining without cloning an extended sequence of a gene specific
CC polynucleotide that has a partial gene sequence known by carrying out the
CC method and then sequencing the amplified isolated gene specific
CC polynucleotide without cloning. The method has several diagnostic and
CC research applications, including genetic and infectious disease
CC diagnosis, forensic and paternity testing, toxicology testing, individual
CC identification, genetic research, occupational hazard screening and
CC pharmaceutical development. The method can be used to detect mutations in
CC target genes which may lead to disease conditions such as Alzheimer's
CC disease, cancer and Duchenne muscular dystrophy, thus facilitating the
CC diagnosis of genetic diseases as well as the identification of carriers
CC of recessive genetic disorders. The present invention is an
CC oligonucleotide useful in the method of the invention
XX
SQ Sequence 43 BP; 10 A; 4 C; 7 G; 22 T; 0 U; 0 Other;
Query Match 70.4%; Score 17.6; DB 6; Length 43;
Best Local Similarity 83.3%; Pred. No. 2.8e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 AAAAAAATCCTGTAGGCTTTTCAGT 24
Db 31 AAAAAAATCCTGTAGGCTTTTCAGT 8
RESULT 43
ADB50438/c
ID ADB50438 standard; DNA; 457 BP.
XX ADB50438;
XX 04-DEC-2003 (first entry)
XX Primary rat hepatocyte toxicity modelling related gene SEQ ID NO:980.
XX toxic effect; gene expression profile; hepatotoxicity; diagnostic marker;
XX toxicity marker; toxicity progression; drug screening;
XX primary rat hepatocyte toxicity modelling; gene; ds.
XX Rattus norvegicus.
XX W02003065993-A2.
XX 14-AUG-2003.
XX 04-FEB-2003; 2003WO-US003482.
XX 04-FEB-2002; 2002US-0353171P.
XX 13-MAR-2002; 2002US-0363534P.
PR

PR 08-APR-2002; 2002US-0370248P.
PR 10-APR-2002; 2002US-0371134P.
PR 10-APR-2002; 2002US-0371135P.
PR 10-APR-2002; 2002US-0371150P.
PR 11-APR-2002; 2002US-0371413P.
PR 19-APR-2002; 2002US-0373601P.
PR 19-APR-2002; 2002US-0373602P.
PR 22-APR-2002; 2002US-0374139P.
PR 08-MAY-2002; 2002US-0378370P.
PR 09-MAY-2002; 2002US-0378652P.
PR 09-MAY-2002; 2002US-0378653P.
PR 09-MAY-2002; 2002US-0378665P.
PR 09-JUL-2002; 2002US-0394230P.
PR 09-JUL-2002; 2002US-0394253P.
PR 04-SEP-2002; 2002US-0407688P.
PR 28-JAN-2003; 2003US-0442900P.
XX (GENE-) GENE LOGIC INC.
PA Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Orr M;
XX Blashoff M;
PI WPI; 2003-731472/69.
XX Determining if a compound induces a toxic effect on a tissue or cell, for
XX identifying hepatotoxic compounds, comprises comparing a gene expression
PT profile of a tissue or cell sample to a database of Tox mean and non-Tox
PT mean values.
XX Claim 44; SEQ ID NO 980; 874pp; English.
XX The present invention describes a method for determining whether a
XX compound induces a toxic effect on a tissue or cell. The method comprises
CC preparing a gene expression profile of a tissue or cell sample exposed to
CC the compound, and comparing the gene expression profile to a database
CC comprising data or information on the Tox mean and non-Tox mean value.
CC The method is useful for predicting or identifying at least one toxic
CC effect, particularly hepatotoxicity, of a test or unknown compound. The
CC genes listed in the specification are useful as diagnostic or toxicity
CC markers for the prediction or identification of the physiological state
CC of tissue or cell sample that has been exposed to a compound, or to
CC identify or predict the toxic effects of a compound or an agent. These
CC may also be used as markers for monitoring toxicity progression or for
CC drug screening. The present sequence represents a primary rat hepatocyte
CC toxicity modelling related gene sequence from the present invention.
XX
SQ Sequence 457 BP; 139 A; 103 C; 102 G; 113 T; 0 U; 0 Other;
Query Match 70.4%; Score 17.6; DB 10; Length 457;
Best Local Similarity 83.3%; Pred. No. 4e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 AAAAAAATCCTGTAGGCTTTTCAGT 24
Db 262 AAAAAAATCCTGTAGGCTTTTCAGT 239
RESULT 44
ADS56904/c
ID ADS56904 standard; cDNA; 925 BP.
XX ADS56904;
XX 02-DEC-2004 (first entry)
XX Bacterial polynucleotide #8891.
XX Recombinant DNA construct; transformed plant; improved plant property;
XX cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
XX pathogen tolerance; pest tolerance; plant disease resistance;
XX cell cycle pathway modification; plant growth regulator;
XX homologous recombination; seed oil yield; protein yield; carbohydrate;
XX nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
XX

```

KW bacterial polynucleotide; gene; ss.
XX
OS Bacteria.
XX
XX US2003233675-A1.
XX
PD 18-DEC-2003.
XX
XX 20-FEB-2003; 2003US-00369493.
XX
XX 21-FEB-2002; 2002US-0360039P.
XX
XX (CAOY/) CAO Y.
XX (HINK/) HINKLE G J.
XX (SLAT/) SLATER S C.
XX (CHEN/) CHEN X.
XX (GOLD/) GOLDMAN B S.
XX
XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX
XX WPI; 2004-061375/06.
XX
XX New recombinant DNA construct comprising a promoter positioned to provide
XX for expression of a polynucleotide encoding a polypeptide from a
XX microbial source, useful for producing plants with improved properties.
XX
XX Claim 1; SEQ ID NO 32578; 122pp; English.
XX
XX The invention relates to a recombinant DNA construct comprising a
XX promoter functional in a plant cell, where the promoter is positioned to
XX provide for expression of a polynucleotide encoding a polypeptide from a
XX microbial source. The invention also relates to a transformed plant
XX comprising the recombinant DNA construct and a method of producing a
XX transformed plant having an improved property. The plant is a crop plant
XX such as maize or soybean. The method of producing a transformed plant
XX having an improved property comprises transforming a plant with the
XX recombinant DNA construct and growing the transformed plant, where the
XX polynucleotide or polypeptide is useful for improving plant properties.
XX The recombinant DNA construct is useful for producing plants with
XX improved plant properties, e.g. improved cold, heat or drought tolerance,
XX tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
XX increased resistance to plant disease, better growth rate by modification
XX of the cell cycle pathway with plant growth regulators, increased rate of
XX homologous recombination, modified seed oil or protein yield and/or
XX content, improved yield by modification of carbohydrate, nitrogen or
XX phosphorus use and/or uptake, by modification of photosynthesis or by
XX providing improved plant growth and development under at least one stress
XX condition, improved lignin production or improved galactomannan
XX production. This sequence represents a bacterial polynucleotide used in
XX the scope of the invention. Note: The sequence data for this patent did
XX not form part of the printed specification but was obtained in electronic
XX format from USPTO at seqdata.uspto.gov/sequence.html.
XX
XX Sequence 925 BP; 168 A; 237 C; 278 G; 242 T; 0 U; 0 Other;
XX
XX Query Match 70.4%; Score 17.6; DB 13; Length 925;
XX Best Local Similarity 83.3%; Pred. No. 4.5e+02;
XX Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
XX
XX QY 2 AAAAAGCTCTGAGGCTTTTCAGTG 25
XX ||||| ||||| ||||| ||||| |||||
XX 912 AAAAAGCTCTGAGGCTTTTCAGTG 889
XX
XX Db
XX
XX RESULT 45
XX ADX09691/C
XX ID ADX09691 standard; cDNA; 1303 BP.
XX
XX AC ADX09691;
XX
XX XX
XX 21-APR-2005 (first entry)
XX
XX DE Plant full length insert polynucleotide seqid 4266.
XX

XX plant protectant; plant growth regulant; gene therapy; plant;
KW recombinant DNA construct; physical array; plant breeding marker;
KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
KW extreme osmotic condition; pathogen tolerance; pest tolerance;
KW growth rate; cell cycle pathway; disease resistance;
KW galactomannan production; lignin production; plant growth regulator;
KW yield; plant growth; plant development; seed oil; protein yield;
XX protein content; gene; ss.
XX
XX Unidentified.
OS
XX
XX US2004034888-A1.
XX
XX 19-FEB-2004.
PD
XX
XX 28-APR-2003; 2003US-00425114.
XX
XX 06-MAY-1999; 99US-00304517.
XX 05-NOV-2001; 2001US-00985678.
XX
XX (LIUJ/) LIU J.
XX (ZHOU/) ZHOU Y.
XX (KOVA/) KOVALIC D K.
XX (SCRE/) SCREEN S E.
XX (TABA/) TABASKA J E.
XX (CAOY/) CAO Y.
XX
XX Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
XX
XX WPI; 2004-180133/17.
XX
XX New recombinant DNA construct, useful for improving plant tolerance to
XX cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
XX pests, for conferring increased resistance to plant disease, or for
XX improving yield.
XX
XX Claim 1; SEQ ID NO 4266; 15pp; English.
XX
XX The invention describes a recombinant DNA construct comprising a
XX polynucleotide consisting of a sequence encoding an amino acid sequence
XX available in electronic form from the US patent office at
XX ftp.seqdata.uspto.gov/sequence.html?docID:2004034888. The polynucleotide
XX of the invention are also useful in physical arrays of molecules and as
XX plant breeding markers. The recombinant DNA construct is useful for
XX improving plant tolerance to cold, heat, drought, herbicides, extreme
XX osmotic conditions, pathogens or pests, for manipulating growth rate in
XX plant cells by modification of the cell cycle pathway, for conferring
XX increased resistance to plant disease, for producing galactomannan,
XX lignin or plant growth regulators, for increasing the rate of homologous
XX recombination in plants, for improving yield by modification of
XX photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
XX or by providing improved plant growth and development under at least one
XX stress condition or for modifying seed oil or protein yield and/or
XX content. This sequence represents a plant full length insert
XX polynucleotide that can be used in the recombinant DNA construct of the
XX invention.
XX
XX Sequence 1303 BP; 381 A; 278 C; 257 G; 387 T; 0 U; 0 Other;
XX
XX Query Match 70.4%; Score 17.6; DB 13; Length 1303;
XX Best Local Similarity 83.3%; Pred. No. 4.8e+02;
XX Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
XX
XX QY 1 AAAAAGCTCTGAGGCTTTTCAGT 24
XX ||||| ||||| ||||| ||||| |||||
XX 208 AAAAAGCTCTGAGGCTTTTCAGT 185
XX
XX Db
XX
XX RESULT 46
XX ABL11811
XX ID ABL11811 standard; cDNA; 1368 BP.
XX
XX
```


PA (SPIN/) SPINDEL E R.
PA (NAGA/) NAGALLA S.
PA (BARR/) BARRY B.
XX
PI Spindel ER, Nagalla S, Barry B;
XX
DR WPI: 2001-488385/53.
DR P-PSDB; AAU07635.
XX
XX New bombesin like peptide with high binding affinity for bombesin
PT receptors for the modulation of receptors, for example, in the
PT gastrointestinal tract.
XX
PS Claim 15; Fig 1; 16pp; English.
XX
XX The invention relates to genes which encode receptors capable of binding
CC to bombesin-like peptides, expressed in vertebrate tissue. Bombesin-like
CC peptides are produced and secreted by human small cell lung cancer cells.
CC Examples of these peptides are gastrin releasing peptide (GRP),
CC neuromedin B (NMB), neuromedin C (NMC), bombesin receptor subtype-3 (BRS-
CC 3), BB4 and BB5. The peptides of the invention are useful for modulating
CC receptors to which they have an affinity including receptors in the
CC gastrointestinal tract, the central nervous system, testes and uterus.
CC The sequences are therefore used in the treatment of cancers such as
CC small cell lung cancer. This sequence represents a cDNA encoding the frog
CC BB4 receptor
XX
SQ Sequence 1563 BP; 450 A; 333 C; 296 G; 484 T; 0 U; 0 Other;
Query Match 70.4%; Score 17.6; DB 5; Length 1563;
Best Local Similarity 83.3%; Pred. No. 4.9e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 AAAAAAAGCTCTAGGCTTTTCAGT 24
DB 187 AAAAAATGCTGAAGTTTTCAGT 210
RESULT 49
ADE83485/c
ID ADE83485 standard; DNA; 2962 BP.
XX
AC ADE83485;
XX
XX 29-JAN-2004 (first entry)
XX Human gene NM_005419, SEQ ID NO 11081.
XX
XX Human; ds; gene; pain; neuronal tissue; gene therapy;
KW spinal segmental nerve injury; chronic constriction injury; CCI;
KW spared nerve injury; SNI; Chung.
XX
XX Homo sapiens.
XX WO2003016475-A2.
XX
XX 27-FEB-2003.
XX
XX 14-AUG-2002; 2002WO-US025765.
XX
XX 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333347P.
XX
XX (GEHO) GEN HOSPITAL CORP.
PA (FARB) BAYER AG.
XX
XX Woolf C, D'urso D, Befort K, Costigan M;
XX
XX WPI: 2003-268312/26.
DR GENBANK; NM_005419.
XX
XX New composition comprising two or more isolated polypeptides, useful for

PT preparing a medicament for treating pain in an animal.
XX
PS Claim 1; Page; 1017pp; English.
XX
XX The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regularly
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a human DNA (shown in Table 2 of the
CC specification) which encodes one of the polypeptides of the invention
CC which is differentially expressed during pain. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 2962 BP; 749 A; 763 C; 758 G; 692 T; 0 U; 0 Other;
Query Match 70.4%; Score 17.6; DB 10; Length 2962;
Best Local Similarity 83.3%; Pred. No. 5.4e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 AAAAAAAGCTCTAGGCTTTTCAGT 24
DB 520 AGAAGACTCTGCTGCTTTTCAGT 497
RESULT 50
ABLL1810/c
ID ABLL1810 standard; cDNA; 5010 BP.
XX
AC ABLL1810;
XX
XX 26-MAR-2002 (first entry)
XX
XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 29912.
XX
XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ss.
XX
XX Drosophila melanogaster.
XX
XX WO200171042-A2.
XX
XX 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US009231.
XX
XX 23-MAR-2000; 2000US-0191637P.
PR 11-JUL-2000; 2000US-00614150.
XX
XX (PEKE) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
XX
XX WPI; 2001-656860/75.
DR

DR P-PSDB; ABB67707.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX
XX
PS Claim 1; SEQ ID NO 29912; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 5010 BP; 1462 A; 1029 C; 1048 G; 1471 T; 0 U; 0 Other;
Query Match 70.4%; Score 17.6; DB 4; Length 5010;
Best Local Similarity 83.3%; Pred. No. 5.9e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 AAAAACTCCTGTAGGCTTTCAGT 24
Db 1957 AGAAAACTCCTGTGGGCTATCAGT 1934

Search completed: February 3, 2006, 21:56:58
Job time : 212.111 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 3, 2006, 14:38:25 ; Search time 2951.67 Seconds
(without alignments)
396.277 Million cell updates/sec

Title: US-10-719-900-20

Perfect score: 25

Sequence: 1 aaaaactcctgtaggcttctcagt 25

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database :

EST:*

1: gb_est1.*

2: gb_est2.*

3: gb_est3.*

4: gb_est4.*

5: gb_est5.*

6: gb_est6.*

7: gb_est7.*

8: gb_est8.*

9: gb_est9.*

10: gb_est10.*

11: gb_est11.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	25	100.0	706	2	BB474887 BB474887
2	25	100.0	3793	4	AK084510 Mus muscu
3	25	100.0	4437	4	AK036338 Mus muscu
4	20.2	80.8	697	2	BF218988 601882209
5	19.8	79.2	634	8	DN755785 GL-Cf-125
6	19.8	79.2	704	8	DN267760 LI30320-
7	19.8	79.2	715	8	DN265405 LI30320-
8	19.8	79.2	765	8	DN392910 LI3934-0
9	19.8	79.2	793	8	DN370580 LI3733-0
10	19.8	79.2	892	8	CV804204 AGENCOURT
11	19.8	79.2	970	6	CD502134 CDA55-A01
12	19.4	77.6	750	5	BX921863 BX921863
13	19.4	77.6	961	11	AL284798 Tetraodon
14	19.2	76.8	154	1	AI918165 t08a05.x
15	19.2	76.8	163	6	CD544491 B0258C10-
16	19.2	76.8	169	3	BM769029 K-EST0052
17	19.2	76.8	186	1	AA864560 oh31b01.s
18	19.2	76.8	203	1	AI886941 w194C12.x
19	19.2	76.8	226	1	AV232251 AV232251
20	19.2	76.8	226	3	BI687636 603312626
21	19.2	76.8	236	1	AV255052 AV255052
22	19.2	76.8	239	5	BY372216 BY372216

pBluescript KS(+) after bulk excision from Lambda FLC I."

ORIGIN		TITLE	
Query Match	100.0%; Score 25; DB 2; Length 706;	Direct Submission	
Best Local Similarity	100.0%; Pred. No. 1.3;	Submitted (16-APR-2002)	Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
Matches	25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.	
QY	1 AAAAAAATCTCTAGGCTTTTCAGTG 25 	URL: http://genome.gsc.riken.jp/ URL: http://phantom.gsc.riken.jp/ Location/Qualifiers	
Db	575 AAAAAAATCTCTAGGCTTTTCAGTG 599 	1. .3793 /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6J" /db_xref="FANTOM DB:D330010D18" /db_xref="taxon:10090" /clone="D330010D18" /tissue type="heart" /clone_lib="RIKEN full-length enriched mouse cDNA library" /dev_stage="13 days embryo" 1067. .2155 /note="unannotated protein product; mesoderm development candidate 1 (MGD GRI:1891420, GB NM_030705, evidence: BLASTN, 99%, match=3258) putative" /codon_start=1 /protein_id="BAC39204.1" /db_xref="GI:26351135" OLVADLLLSSEARPVLFEGPASPAGAESFECRTDIIARTKGLSLTHDVOSQLNM GRFGAGDSLVGLDVLVSLTSCSAHAAYLAAVATPAQAPGLVDYRVTRCHRV EGCAVLRTATPLADMTPTQLLVSQGLSRNLKELTDACALASDKSRDFRSEOFKLG KCMTSASALLACVREVKAAPSRALSRCLFSGPLVQVAVSLGFATEPQFLGRAAA VTEGKAVQTALIGGMSVSVACVLLTQCLRLAQHPDGSKMSDHRRLRNSACAVS EGCTLLSQALRRSSPRTLPPVNSNVN"	
RESULT 2			
AK084510			
LOCUS			
DEFINITION	3793 bp mRNA linear HTC 03-APR-2004 Mus musculus 13 days embryo heart cDNA, RIKEN full-length enriched library, clone:D330010D18 product:mesoderm development candidate 1, full insert sequence.		
ACCESSION	AK084510.1 GI:26351134		
VERSION			
KEYWORDS	HTC; CAP trapper.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE			
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.		
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes		
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)		
PUBMED	11042159		
REFERENCE			
AUTHORS	Carninci, P., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamanoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, B., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.		
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer		
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)		
PUBMED	11075861		
REFERENCE			
AUTHORS	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.		
TITLE	Functional annotation of a full-length mouse cDNA collection		
JOURNAL	Nature 409, 685-690 (2001)		
REFERENCE			
AUTHORS	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.		
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs		
JOURNAL	Nature 420, 563-573 (2002)		
REFERENCE			
AUTHORS	Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayaehida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuura, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ono, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,		
QY	1 AAAAAAATCTCTAGGCTTTTCAGTG 25 	Query Match	100.0%; Score 25; DB 4; Length 3793;
Db	3662 AAAAAAATCTCTAGGCTTTTCAGTG 3686 	Best Local Similarity	100.0%; Pred. No. 1.8;
RESULT 3		Matches	25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
AK036338			
LOCUS			
DEFINITION	4437 bp mRNA linear HTC 03-APR-2004 Mus musculus 16 days neonate cerebellum cDNA, RIKEN full-length enriched library, clone:9630059C21 product:mesoderm development candidate 1, full insert sequence.		
ACCESSION	AK036338.1 GI:26085172		
VERSION			
KEYWORDS	HTC; CAP trapper.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE			
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
TITLE	Carninci, P. and Hayashizaki, Y.		
JOURNAL	High-efficiency full-length cDNA cloning		
PUBMED	10349636		

```

2
REFERENCE
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
PUBMED 11042159
REFERENCE
AUTHORS 3
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsumai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwaki, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
PUBMED 11076861
REFERENCE
AUTHORS 4
The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
REFERENCE
AUTHORS 5
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
REFERENCE
AUTHORS 6 (bases 1 to 4437)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.
TITLE Direct Submission
JOURNAL Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.jp,
URL: http://genome.gsc.riken.jp/, tel: 81-45-503-9222,
Fax: 81-45-503-9216)
COMMENT cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/
FEATURES
source
Location/Qualifiers
1..4437
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/db_xref="FANTOM_DB:9630059C21"
/db_xref="taxon:10090"
/cloned="9630059C21"
/tissue_type="cerebellum"
/cloned_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="16 days neonate"
1086..2175
/notes="mesoderm development candidate 1 (MGD|MG1:1891420,
GB|NM_030705, evidence: BLASTN, 99%, match=3258)

```

```

polyA_signal putative"
4413..4418
/note="putative"
polyA_site 4437
/note="putative"
ORIGIN
Query Match 100.0%; Score 25; DB 4; Length 4437;
Best Local Similarity 100.0%; Pred. No. 1.8; Indels 0; Gaps 0;
Matches 25; Conservative 0; Mismatches 0;
QY 1 AAAAAAATCCTGTAGGCTTTTCAGTG 25
Db 3682 AAAAAAATCCTGTAGGCTTTTCAGTG 3706
RESULT 4
BF218988/c 697 bp mRNA linear EST 06-NOV-2000
LOCUS 601882209F1 NIH_MGC_57 Homo sapiens cDNA clone IMAGE:4094499 5',
DEFINITION mRNA sequence.
ACCESSION BF218988
VERSION BF218988.1 GI:11112575
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
REFERENCE 1 (bases 1 to 697)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: CLONETECH Laboratories, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCW955 row: b column: 04
High quality sequence stop: 613.
FEATURES
source
Location/Qualifiers
1..697
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cloned="IMAGE:4094499"
/tissue_type="glioblastoma"
/lab_host="DH10B (T1 phage-resistant)"
/cloned_lib="NIH MGC 57"
/notes="Organ: brain; Vector: pDNR-LIB (Clontech); Site: 1:
SfiI (ggccgctcgcc); Site 2: SfiI (ggccattggcc);
SfiI (ggccgctcgcc); Site 3: SfiI (ggccattggcc);
Double-stranded cDNA was prepared from cell line RNA.
and 3' adaptors were used in cloning as follows: 5'
adaptor sequence: 5'-CACGGCATATGACC-3' and 3' adaptor
sequence: 5'-ATTCTAGAGCGGAGCGGCGCATG-dt(30)BN-3'
(where B = A, C, or G and N = A, C, G, or T). Average
insert size 1.55 kb (range 0.9-4.0 kb). 12/15 colonies
contained inserts by PCR. This library was enriched for
full-length clones and was constructed by Clontech
Laboratories (Palo Alto, CA)."
ORIGIN
Query Match 80.8%; Score 20.2; DB 2; Length 697;
Best Local Similarity 88.0%; Pred. No. 2.3e+02;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 AAAAAAATCCTGTAGGCTTTTCAGTG 25
Db 557 AAAAAAATCCTGTAGGCTTTTCATTTG 533

```

RESULT 5
DN755785
LOCUS
DEFINITION
GL-CF-12581 GLGC-LIB0001-cf Canis familiaris Normalized Mixed
Tissue cDNA Library Canis familiaris cDNA, mRNA sequence.

ACCESSION
DN755785
VERSION
DN755785.1 GI:62144898
KEYWORDS
EST.
SOURCE
Canis familiaris (dog)
ORGANISM
Canis familiaris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;
Canis.

REFERENCE
1 (bases 1 to 634)
Liu, Q.
Direct Submission Gene Logic Inc
TITLE
Unpublished (2005)
JOURNAL
Contact: Qing Liu
COMMENT
Gene Logic Inc.
610 Professional Drive, Gaithersburg, MD 20879, USA
Tel: 301 987 1700
Email: qlu@genelogic.com.

FEATURES
source
1..634
Location/Qualifiers

/organism="Canis familiaris"
/mol_type="mRNA"
/strain="beagle"
/db_xref="taxon:9615"
/lab_host="EMDH10B"
/clone_lib="GLGC-LIB0001-cf Canis familiaris Normalized
Mixed Tissue cDNA Library"
note="Organ: heart, liver, kidney, testis, and brain;
Vector: pCMVSPORT6.0; Site_1: NotI; Site_2: SalI"

ORIGIN

Query Match 79.2%; Score 19.8; DB 8; Length 634;
Best Local Similarity 91.3%; Pred. No. 3.5e+02;
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AAAAATCCTGTAGGCTTCAGT 24
||||| ||||| ||||| ||||| |||||

Db 313 AAAAATCCTGTAGGCTTCAGT 335

RESULT 6
DN267760
LOCUS
DEFINITION
LIB30320_023 G09 SP6.1 LIB30320 Canis familiaris cDNA clone
LIB30320_023_G09, mRNA sequence.

ACCESSION
DN267760
VERSION
DN267760.1 GI:60436370
KEYWORDS
EST.
SOURCE
Canis familiaris (dog)

ORGANISM
Canis familiaris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;
Canis.

REFERENCE
1 (bases 1 to 704)
Staten, N.R.
Direct Submission (Staten, N.R.)
TITLE
Unpublished (2005)
JOURNAL
Contact: Nick Staten

COMMENT
Tel: 636 247 6855
Email: nicholas.r.staten@pfizer.com.

FEATURES
source
1..704
Location/Qualifiers
/organism="Canis familiaris"
/mol_type="mRNA"
/db_xref="taxon:9615"
/clone="LIB30320_023_G09"
/tissue_type="lymph node"

ORIGIN

Query Match 79.2%; Score 19.8; DB 8; Length 704;
Best Local Similarity 91.3%; Pred. No. 3.5e+02;
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AAAAATCCTGTAGGCTTCAGT 24
||||| ||||| ||||| ||||| |||||

Db 306 AAAAATCCTGTAGGCTTCAGT 328

RESULT 7
DN265405
LOCUS
DEFINITION
LIB30320_035_A07 SP6.1 LIB30320 Canis familiaris cDNA clone
LIB30320_035_A07, mRNA sequence.

ACCESSION
DN265405
VERSION
DN265405.1 GI:60434015
KEYWORDS
EST.
SOURCE
Canis familiaris (dog)

ORGANISM
Canis familiaris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;
Canis.

REFERENCE
1 (bases 1 to 715)
Staten, N.R.

AUTHORS
Direct Submission (Staten, N.R.)
TITLE
Unpublished (2005)
JOURNAL
Contact: Nick Staten

COMMENT
Tel: 636 247 6855
Email: nicholas.r.staten@pfizer.com.

FEATURES

source
1..715
Location/Qualifiers
/organism="Canis familiaris"
/mol_type="mRNA"
/db_xref="taxon:9615"
/clone="LIB30320_035_A07"
/tissue_type="lymph node"
/lab_host="EMDH10B"
/clone_lib="LIB30320"
/note="Vector: pCMV-Sport6; Site_1: NotI; Site_2: SalI;
barcode 459"

ORIGIN

Query Match 79.2%; Score 19.8; DB 8; Length 715;
Best Local Similarity 91.3%; Pred. No. 3.5e+02;
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AAAAATCCTGTAGGCTTCAGT 24
||||| ||||| ||||| ||||| |||||

Db 346 AAAAATCCTGTAGGCTTCAGT 368

RESULT 8
DN392910
LOCUS
DEFINITION
LIB3934-034-A1-K1-F8 LIB3934 Canis familiaris cDNA clone
CLN12935017, mRNA sequence.

ACCESSION
DN392910
VERSION
DN392910.1 GI:60574131
KEYWORDS
EST.

SOURCE

ORGANISM
Canis familiaris (dog)
Canis familiaris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;
Canis.

REFERENCE
1 (bases 1 to 765)
Staten, N.R.

AUTHORS
Direct Submission (Staten, N.R.)
TITLE

SOURCE ORGANISM
 Gasterosteus aculeatus (three spined stickleback)
 Gasterosteus aculeatus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 Acanthomorpha; Acanthopterygii; Percomorpha; Gasterosteiformes;
 Gasterosteidae; Gasterosteus.
 1 (bases 1 to 970)

REFERENCE AUTHORS
 Kingsley, D.M., Peichel, C., Balabhadra, S., Grimwood, J., Dickson, M.,
 Schmutz, J. and Myers, R.M.
 Expressed sequence tags from Gasterosteus aculeatus
 Unpublished (2003)
 Contact: Kingsley, DM
 HHMI and Department of Developmental Biology
 Stanford University School of Medicine
 Beckman Center B300, 279 Campus Drive, Stanford, CA 94305-5329, USA
 Tel: 650 725 5954
 Fax: 650 725 7739
 Email: kingsley@cngm.stanford.edu
 Plate: 55
 High quality sequence start: 5
 High quality sequence stop: 837.
 Location/Qualifiers
 1..970

FEATURES
 source
 /organism="Gasterosteus aculeatus"
 /mol_type="mRNA"
 /strain="Salinas river, CA"
 /db_xref="taxon:69293"
 /clone="CDAS5-A01"
 /sex="mixed male and female"
 /tissue_type="heads and internal organs combined"
 /dev_stage="adult"
 /clone_lib="SHGC-CDA"
 /note="Vector: lambda ZAP Express/pBK-CMV; Site 1: EcoRI
 (5' adaptor); Site 2: XhoI (3' linker primer); The mixed
 organ cDNA library was generated using the ZAP-cDNA method
 by Stratagene. First strand cDNA synthesis was primed with
 a 50 bp linker primer containing an oligo dT sequence
 preceded by a synthetic XhoI site. 5 prime adaptors were
 used containing an EcoRI cohesive end. The finished cDNAs
 were inserted in to the ZAP express vector
 unidirectionally in the sense orientation with respect to
 the lacZ promoter of pBK-CMV. An amplified library was
 prepared from approximately 3 million primary clones in
 the lambda ZAP Express vector. In vivo excision was then
 used to generate individual pBK-CMV phagemid clones for
 EST sequencing."

ORIGIN
 Query Match 79.2%; Score 19.8; DB 6; Length 970;
 Best Local Similarity 91.3%; Pred. No. 3.8e+02;
 Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 2 AAAAACTCTGTAGGCTTCAGT 24
 ||||| ||||| ||||| |||||
 Db 38 AAAAACACCTGGAGGCTTCAGT 60

RESULT 12
 BX921863/c
 LOCUS
 DEFINITION
 BX921863 Sus Scrofa library (scan) Sus scrofa cDNA clone
 scan0019d.o.08 5prim, mRNA sequence.
 BX921863
 BX921863.1 GI:41138711
 EST.
 Sus scrofa (pig)
 Sus scrofa
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae;
 Sus.
 1 (bases 1 to 750)
 Bonnet, A., Tosser-Klopp, G., Benne, F., Cabau, C., Villegier, S.,
 Soares, M., Bonaldo, F. and Hately, F.

SOURCE ORGANISM
 A Pig Normalised Multi-Tissue cDNA Library
 Unpublished (2003)
 Contact: Tosser-Klopp G
 Genetique Animale
 Institut National de la Recherche Agronomique
 Chemin de Borde-Rouge - Auzeville BP27, 31326 Castanet-Tolosan
 cedex, FRANCE
 Tel: 33 (0) 5.61.28.51.14
 Fax: 33 (0) 5.61.28.53.08
 Email: tosser@toulouse.inra.fr
 Sequence cleaned of vector adaptor and repetitions. Contact us
 at aigenasupport@jouy.inra.fr to obtain the Chromatogram of this
 sequence.
 Plate: 0019 row: 0 column: 8.
 Location/Qualifiers
 1..750
 /organism="Sus scrofa"
 /mol_type="mRNA"
 /db_xref="taxon:9823"
 /clone="scan0019d.o.08"
 /tissue_type="mixed"
 /dev_stage="from embryos to adults"
 /clone_lib="Sus Scrofa library (scan)"
 /note="tissues: adipose tissue, brain, kidney, liver,
 muscle, ovary, testis, heart, hypothalamus, pancreas,
 skin, spleen, thymus, placenta, pituitary gland, seminal
 vesicle, small intestine, uterus, adrenals, bulbo urethral
 gland, cerebral trunk, epididymis, female gonad,
 gall-bladder, hippocampus, large intestine, male gonad,
 melanocytes, stomach, udder"

ORIGIN
 Query Match 77.6%; Score 19.4; DB 5; Length 750;
 Best Local Similarity 95.2%; Pred. No. 5.5e+02;
 Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 AAAAACTCTGTAGGCTTTC 21
 ||||| ||||| ||||| |||||
 Db 412 AAAAAGTCCTGTAGGCTTTC 392

RESULT 13
 CNS04CQT/c
 LOCUS
 DEFINITION
 CNS04CQT 961 bp DNA linear GSS 01-SEP-2000
 Tetraodon nigroviridis genome survey sequence T7 end of clone
 101C23 of library G from Tetraodon nigroviridis, genomic survey
 sequence.
 AL284798
 AL284798.1 GI:8023197
 GSS; genome survey sequence.
 Tetraodon nigroviridis
 Tetraodon nigroviridis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
 Tetraodontoidea; Tetraodontidae; Tetraodon.
 1
 Roest Crolius, H., Jaillon, O., Dasilva, C., Bouneau, L., Fisher, C.,
 Bernot, A., Fizames, C., Wincker, P., Brottier, P., Quetier, F.,
 Saurin, W. and Weissenbach, J.
 Estimate of human gene number provided by genome-wide analysis
 using Tetraodon nigroviridis DNA sequence
 Nat. Genet. 25 (2), 235-238 (2000)
 10835645
 2
 Roest Crolius, H., Jaillon, O., Dasilva, C., Ozouf-Costaz, C.,
 Fizames, C., Fischer, C., Bouneau, L., Billault, A., Quetier, F.,
 Saurin, W., Bernot, A. and Weissenbach, J.
 Characterization and repeat analysis of the compact genome of the
 freshwater pufferfish Tetraodon nigroviridis
 Genome Res. 10 (7), 939-949 (2000)
 10899143
 3 (bases 1 to 961)

TITLE JOURNAL COMMENT
 A Pig Normalised Multi-Tissue cDNA Library
 Unpublished (2003)
 Contact: Tosser-Klopp G
 Genetique Animale
 Institut National de la Recherche Agronomique
 Chemin de Borde-Rouge - Auzeville BP27, 31326 Castanet-Tolosan
 cedex, FRANCE
 Tel: 33 (0) 5.61.28.51.14
 Fax: 33 (0) 5.61.28.53.08
 Email: tosser@toulouse.inra.fr
 Sequence cleaned of vector adaptor and repetitions. Contact us
 at aigenasupport@jouy.inra.fr to obtain the Chromatogram of this
 sequence.
 Plate: 0019 row: 0 column: 8.
 Location/Qualifiers
 1..750
 /organism="Sus scrofa"
 /mol_type="mRNA"
 /db_xref="taxon:9823"
 /clone="scan0019d.o.08"
 /tissue_type="mixed"
 /dev_stage="from embryos to adults"
 /clone_lib="Sus Scrofa library (scan)"
 /note="tissues: adipose tissue, brain, kidney, liver,
 muscle, ovary, testis, heart, hypothalamus, pancreas,
 skin, spleen, thymus, placenta, pituitary gland, seminal
 vesicle, small intestine, uterus, adrenals, bulbo urethral
 gland, cerebral trunk, epididymis, female gonad,
 gall-bladder, hippocampus, large intestine, male gonad,
 melanocytes, stomach, udder"

FEATURES
 source
 /organism="Sus scrofa"
 /mol_type="mRNA"
 /db_xref="taxon:9823"
 /clone="scan0019d.o.08"
 /tissue_type="mixed"
 /dev_stage="from embryos to adults"
 /clone_lib="Sus Scrofa library (scan)"
 /note="tissues: adipose tissue, brain, kidney, liver,
 muscle, ovary, testis, heart, hypothalamus, pancreas,
 skin, spleen, thymus, placenta, pituitary gland, seminal
 vesicle, small intestine, uterus, adrenals, bulbo urethral
 gland, cerebral trunk, epididymis, female gonad,
 gall-bladder, hippocampus, large intestine, male gonad,
 melanocytes, stomach, udder"

ORIGIN
 Query Match 77.6%; Score 19.4; DB 5; Length 750;
 Best Local Similarity 95.2%; Pred. No. 5.5e+02;
 Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 AAAAACTCTGTAGGCTTTC 21
 ||||| ||||| ||||| |||||
 Db 412 AAAAAGTCCTGTAGGCTTTC 392

RESULT 13
 CNS04CQT/c
 LOCUS
 DEFINITION
 CNS04CQT 961 bp DNA linear GSS 01-SEP-2000
 Tetraodon nigroviridis genome survey sequence T7 end of clone
 101C23 of library G from Tetraodon nigroviridis, genomic survey
 sequence.
 AL284798
 AL284798.1 GI:8023197
 GSS; genome survey sequence.
 Tetraodon nigroviridis
 Tetraodon nigroviridis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
 Tetraodontoidea; Tetraodontidae; Tetraodon.
 1
 Roest Crolius, H., Jaillon, O., Dasilva, C., Bouneau, L., Fisher, C.,
 Bernot, A., Fizames, C., Wincker, P., Brottier, P., Quetier, F.,
 Saurin, W. and Weissenbach, J.
 Estimate of human gene number provided by genome-wide analysis
 using Tetraodon nigroviridis DNA sequence
 Nat. Genet. 25 (2), 235-238 (2000)
 10835645
 2
 Roest Crolius, H., Jaillon, O., Dasilva, C., Ozouf-Costaz, C.,
 Fizames, C., Fischer, C., Bouneau, L., Billault, A., Quetier, F.,
 Saurin, W., Bernot, A. and Weissenbach, J.
 Characterization and repeat analysis of the compact genome of the
 freshwater pufferfish Tetraodon nigroviridis
 Genome Res. 10 (7), 939-949 (2000)
 10899143
 3 (bases 1 to 961)

TITLE JOURNAL COMMENT
 A Pig Normalised Multi-Tissue cDNA Library
 Unpublished (2003)
 Contact: Tosser-Klopp G
 Genetique Animale
 Institut National de la Recherche Agronomique
 Chemin de Borde-Rouge - Auzeville BP27, 31326 Castanet-Tolosan
 cedex, FRANCE
 Tel: 33 (0) 5.61.28.51.14
 Fax: 33 (0) 5.61.28.53.08
 Email: tosser@toulouse.inra.fr
 Sequence cleaned of vector adaptor and repetitions. Contact us
 at aigenasupport@jouy.inra.fr to obtain the Chromatogram of this
 sequence.
 Plate: 0019 row: 0 column: 8.
 Location/Qualifiers
 1..750
 /organism="Sus scrofa"
 /mol_type="mRNA"
 /db_xref="taxon:9823"
 /clone="scan0019d.o.08"
 /tissue_type="mixed"
 /dev_stage="from embryos to adults"
 /clone_lib="Sus Scrofa library (scan)"
 /note="tissues: adipose tissue, brain, kidney, liver,
 muscle, ovary, testis, heart, hypothalamus, pancreas,
 skin, spleen, thymus, placenta, pituitary gland, seminal
 vesicle, small intestine, uterus, adrenals, bulbo urethral
 gland, cerebral trunk, epididymis, female gonad,
 gall-bladder, hippocampus, large intestine, male gonad,
 melanocytes, stomach, udder"

AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source
ORIGIN
Query Match
Best Local Similarity
Matches
QY
Db
RESULT 15
CD544491
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
PUBMED
COMMENT
FEATURES
source

Genoscope.
Direct Submission
Submitted (12-APR-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
This sequence is a single read and was generated as part of a large
scale clone-end sequencing project of the Tetraodon nigroviridis
genome. For more information, please take a look at
http://www.genoscope.cns.fr/Tetraodon.
Location/Qualifiers
1. .961
/organism="Tetraodon nigroviridis"
/mol_type="genomic DNA"
/db_xref="taxon:99883"
/clone="101C23"
/clone_lib="G"
/note="Genoscope sequence ID : COBG101AB12LP1
end : T7"

Query Match 77.6%; Score 19.4; DB 11; Length 961;
Best Local Similarity 95.2%; Pred. No. 5.8e+02;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 AAAACTCCGTAGGCTTTCAG 23
|||||
462 AAAACTCCGTAGGCTTTCG 442

RESULT 14
AI918165
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

AI918165 154 bp mRNA linear EST 13-DEC-1999
tn08a05.x1 NCI_CGAP_Brn25 Homo sapiens cDNA clone IMAGE:2166992 3',
mRNA sequence.
AI918165
AI918165.1 GI:5638020
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 154)
NCI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute / National Institute of Neurological
Disorders and Stroke, Brain Tumor Genome Anatomy Project
(CGAP/BTGP), Tumor Gene Index
Unpublished (1998)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D.,
Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone Distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 1243 Std Error: 0.00
Seq primer: -400P from Gibco
High quality sequence stop: 134.
Location/Qualifiers
1. .154
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2166992"
/tissue_type="anaplastic oligodendroglioma"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Brn25"
/note="Organ: brain; Vector: pT7T3D-Pac (Pharmacia) with a
modified polylinker; Site_1: Not 1; Site_2: Eco RI; 1st

strand cDNA was primed with a Not I - oligo(dT) primer [5'
TGTTACCAATCTGAAGTCGAGCGCCGCATAGGTTTTTTTTTTTTTTTTTTTTT
T 3']; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pT7T3 vector.
Library is normalized, and was constructed by Bento
Soares and M.Fatima Bonaldo."

Query Match 76.8%; Score 19.2; DB 1; Length 154;
Best Local Similarity 87.5%; Pred. No. 5e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAATCTCTGTAGGCTTTCAGT 24
|||||
130 AAAAATCTCTGTAGGCTTTCATT 153

RESULT 15
CD544491
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
PUBMED
COMMENT
FEATURES
source

CD544491 163 bp mRNA linear EST 11-JUN-2003
B0258C10-3 NIA Mouse Embryonic Germ Cell cDNA Library (Long) Mus
musculus cDNA clone NIA:B0258C10 IMAGE:30108993 3', mRNA sequence.
CD544491
CD544491.1 GI:31592226
EST.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Murinae; Mus.
1 (bases 1 to 163)
Piao, Y., Ko, N.T., Lim, M.K. and Ko, M.S.H.
Construction of long-transcript enriched cDNA libraries from
submicrogram amounts of total RNAs by a universal PCR amplification
method
Genome Res. 11 (9), 1553-1558 (2001)
11544199
Contact: Dawood B. Dudekula
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@igsun.grc.nia.nih.gov
Plate: B0258 row: C column: 10
Seq primer: -21M3 Forward
High quality sequence stop: 163
POLYA=Yes.
Location/Qualifiers
1. .163
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CS7BL/6J"
/db_xref="niaEST:B0258C10-3"
/db_xref="taxon:10090"
/clone="NIA:B0258C10 IMAGE:30108993"
/sex="male"
/dev_stage="embryonic day 8"
/lab_host="DH10B"
/clone_lib="NIA Mouse Embryonic Germ Cell cDNA Library
(Long)"
/note="Vector: pCMV-SPORT6 (Invitrogen); Site 1: SalI;
Site 2: NotI; Mouse cDNA project by the Laboratory of
Genetics, National Institute on Aging (NIA), Intramural
Research Program, NIH (http://igsun.grc.nia.nih.gov/cDNA).
This is a long-transcript enriched cDNA library (Ref.
Genome Res. 11: 1553-1558 (2001). [PMID: 11544199]). Total
RNAs were obtained from Dr. Mark G. Carter (NIH/NIA-IRP).
EG cells were cultured at 37 C, 5% CO2 in DMEM
supplemented with 15% ES cell-qualified FBS, 0.1mM
non-essential amino acids, 2 mM glutamine,
penicillin/streptomycin, 1 mM sodium pyruvate, 0.1 mM
beta-mercaptoethanol, and 1000000 units of LIF per liter.
Double-stranded cDNAs were synthesized with an Oligo(dT)

primer [Invitrogen:
5'-pGACTAGTCTAGATCGGAGCGCGCCCTTTT-3'] from
2.5 ug of total RNA, treated with T4 DNA polymerase, and
purified by ethanol-precipitation. The cDNAs were ligated
to lone-linker Lu-Sal4, purified by phenol/chloroform, and
separated from free linkers by Centricon 100. Then, the
cDNAs were amplified by long-range high fidelity PCR using
Ex Taq polymerase (Takara) with a primer Sal4-S. The
products were purified by phenol/chloroform and Centricon
100. The cDNAs were digested with SalI and NotI enzymes
and cloned into SalI/NotI site of pCMV-SPORT6 plasmid
vector. The DH10B E. coli host was transformed with the
ligation mixture by the standard chemical method. The
average insert size is about 4.0 kb. The library was
constructed by Yulan Piao."

ORIGIN

Query Match 76.8%; Score 19.2; DB 6; Length 163;
Best Local Similarity 87.5%; Pred. No. 5e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AAAAACTCTGTAGGCTTTTCAGT 24

Db 123 AAAAAATTCGTAGGTTTCATT 146

RESULT 16
BM769029/c
LOCUS
DEFINITION BM769029 169 bp mRNA linear EST 04-MAR-2002
K-EST0052192 S14K402 Homo sapiens cDNA clone S14K402-14-B08 5',
mRNA sequence.

ACCESSION BM769029.1 GI:19098644

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE
1 (bases 1 to 169)
Kim, N.S., Hahn, Y., Oh, J.H., Lee, J.Y., Ahn, H.Y., Chu, M.Y., Kim, M.R.,
Oh, K.J., Cheong, J.E., Sohn, H.Y., Kim, J.M., Park, H.S., Kim, S. and
Kim, Y.S.

TITLE 21C Frontier Korean EST Project 2001

JOURNAL

COMMENT

Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea

Tel: +82-42-860-4470

Fax: +82-42-860-4409

Email: yongsung@mail.kribb.re.kr

Plate: 14 row: B column: 08

High quality sequence stop: 169.

FEATURES

source

1..169

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="S14K402-14-B08"

/cell_line="K402"

/lab_host="Top10F"

/clone_lib="S14K402"

/note="Organ: Stomach; Vector: pTZ18RP1; Site 1: EcoRI;
Site 2: NotI; The poly (A)+ RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then decapped
with tobacco acid pyrophosphatase (TAP). The decapped
intact mRNA was ligated with DNA-RNA linker including EcoR

I site by treatment of T4 RNA ligase and the first strand
cDNA was synthesized from oligo dt-selected mRNA by
priming with dt-tailed vector. The dt-tailed vector was
adjusted to have about 60nt. The cDNA vector was

circularized with E. coli DNA ligase after digestion of

EcoRI which site is also included in vector. An RNA strand
converted to a DNA strand by Okayama-Berg method. The
obtained cDNA vectors were used for transformation of
competent cells E. coli Top10F by electroporation method.
The cDNA libraries constructed by this method are
full-length enriched cDNA library."

ORIGIN

Query Match 76.8%; Score 19.2; DB 3; Length 169;
Best Local Similarity 87.5%; Pred. No. 5.1e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AAAAACTCTGTAGGCTTTTCAGT 24

Db 36 AAAAAATTCGTAGGTTTCATT 13

RESULT 17

AA864560

LOCUS

DEFINITION

AA864560.1

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.

Emmert-Buck, M.D., Ph.D.

CDNA Library Preparation: Stratagene, Inc.

CDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Cloned through the I.M.A.G.E. Consortium/LLNL at:

www-bio.llnl.gov/bbrp/image/image.html

Insert length: 1284 Std Error: 0.00

Seq primer: -40m13 fwd. ET from Amersham

High quality sequence stop: 111.

FEATURES

source

1..186

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:1459369"

/sex="mixed"

/tissue_type="kidney tumor"

/lab_host="SOLR (kanamycin resistant)"

/clone_lib="NCI CGAP Kid6"

/note="Organ: kidney; Vector: Bluescript SK-; Site 1:

EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer:

Oligo dt. Pooled kidney tumors. 5' adaptor sequence: 5'

GAATTCGGCACGAG 3' 3' adaptor sequence: 5'

CTCGAGTTTTTTTTTTTTTTT 3' Average insert size: 1.0 kb."

Query Match 76.8%; Score 19.2; DB 1; Length 186;

Best Local Similarity 87.5%; Pred. No. 5.2e+02;

Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AAAAACTCTGTAGGCTTTTCAGT 24

Db 145 AAAAAATTCGTAGGTTTCATT 168

RESULT 18	AI886941	203 bp	mRNA	linear	EST 07-MAR-2000
LOCUS	AI886941	203 bp	mRNA	linear	EST 07-MAR-2000
DEFINITION	wl94c12.x1 NCI_CGAP_Brn25 Homo sapiens cDNA clone IMAGE:2432566 3', mRNA sequence.				
ACCESSION	AI886941				
VERSION	AI886941.1	GI:5592105			
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	1 (bases 1 to 203)				
TITLE	NCI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap . National Cancer Institute / National Institute of Neurological Disorders and Stroke, Brain Tumor Genome Anatomy Project (CGAP/BTGP), Tumor Gene Index				
JOURNAL	Unpublished (1998)				
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-r@mail.nih.gov Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D., Ph.D. cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D. cDNA Library Arrayed by: Greg Lennon, Ph.D. DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: www-bio.lnlnl.gov/bbrp/image/image.html Insert Length: 1045 Std Error: 0.00 Seq primer: -40UP from Gibco High quality sequence stop: 198. Location/Qualifiers 1..203 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone=IMAGE:2432566" /tissue_type="anaplastic oligodendroglioma" /lab_host="DH10B" /clone_lib="NCI CGAP Brn25" /note="Organ: brain; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTCGGAGCGCCGATAGGTTTTTTTTTTTTTTTTTT T 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library is normalized, and was constructed by Bento Soares and M.Fatima Bonaldo."				
FEATURES	source				
ORIGIN	Best Match 76.8%; Score 19.2; DB 1; Length 203; Query Match 87.5%; Pred. No. 5.3e+02; Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;				
QY	1 AAAAACTCCTGTAGGCTTTCAGT 24 				
Db	125 AAAAAATTCCTGTAGGTTTCATT 148 				
RESULT 19	AV232251/c	226 bp	mRNA	linear	EST 15-NOV-2001
LOCUS	AV232251	226 bp	mRNA	linear	EST 15-NOV-2001
DEFINITION	RIKEN full-length enriched, 0 day neonate skin Mus musculus cDNA clone 4632413D01 3', similar to AB003502 mRNA for Guanine Nucleotide Regulatory Protein, mRNA sequence.				
ACCESSION	AV232251				
VERSION	AV232251.1	GI:6184766			
KEYWORDS	EST.				
SOURCE	Mus musculus (house mouse)				

Best Local Similarity 87.5%; Pred. No. 5.4e+02; Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;		ORGANISM	Mus musculus
QY	1 AAAAACTCCTAGGCTTTCAGT 24 	REFERENCE AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridea; Muridae; Murinae; Mus. 1 (bases 1 to 236)
Db	97 AAAAAATTCGTAGGTTTTCATT 74 		Konno,H., Aizawa,K., Akahira,S., Akiyama,J., Carninci,P., Endo,T., Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N., Hirozane,T., Hori,F., Ishii,Y., Ishikawa,T., Itoh,M., Izawa,M., Kadota,K., Kagawa,I., Kai,C., Kawai,J., Kikuchi,N., Kojima,Y., Koya,S., Kusakabe,M., Mateuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H., Okazaki,Y., Owa,C., Ozawa,Y., Saito,H., Sano,M., Sato,K., Shibata,K., Shibata,Y., Shigemoto,Y., Shiraki,T., Sogabe,Y., Sugahara,Y., Suzuki,H., Suzuki,H., Takahashi,F., Tateno,M., Tominaga,N., Tsunoda,Y., Watahiki,A., Watanabe,S., Yamamura,T., Yasunishi,A., Yokota,T., Yoshiki,A., Yoshino,M., Muramatsu,M. and Hayaishizaki,Y. RIKEN Mouse ESTs (Konno,H., et al. 1999) Unpublished (1999)
RESULT 20	BI687636 226 bp mRNA linear EST 18-SEP-2001		
LOCUS	603312626F1 NCI_CGAP_Mam6 Mus musculus cDNA clone IMAGE:5352535 5', mRNA sequence.		
DEFINITION	BI687636		
ACCESSION	BI687636.1 GI:15650264		
VERSION	EST.		
KEYWORDS			
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridea; Muridae; Murinae; Mus. 1 (bases 1 to 226)		
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/.		
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)		
JOURNAL	Unpublished (1999)		
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgabbs@mail.nih.gov Tissue Procurement: Jeffrey Green M.D. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLMI1896 row: d column: 08 High quality sequence stop: 226. Location/Qualifiers 1. .226 /organism="Mus musculus" /mol_type="mRNA" /strain="FVB/N" /db_xref="taxon:10090" /clone="IMAGE:5352535" /sex="female, virgin" /tissue_type="infiltrating ductal carcinoma" /dev_stage="5 months" /lab_host="DH10B" /clone_lib="NCI CGAP Mam6" /notes="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI; Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt. Library constructed by Life Technologies. Investigator providing samples: Jeffrey Green, M.D., NIH"		
FEATURES	source	Location/Qualifiers	
	1. .236	/organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6J" /db_xref="taxon:10090" /clone="4921516B04" /sex="male" /tissue_type="testis" /dev_stage="adult" /lab_host="DH10B" /clone_lib="RIKEN full-length enriched, adult male testis (DH10B)"	
ORIGIN			
Query Match	76.8%; Score 19.2; DB 3; Length 226;		
Best Local Similarity 87.5%; Pred. No. 5.4e+02; Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;			
QY	1 AAAAACTCCTAGGCTTTCAGT 24 		
Db	75 AAAAAATTCGTAGGTTTTCATT 52 		
RESULT 21	AV255052 236 bp mRNA linear EST 04-NOV-1999		
LOCUS	AV255052 RIKEN full-length enriched, adult male testis (DH10B) Mus musculus cDNA clone 4921516B04 3' similar to AB003502 Mus musculus mRNA for Guanine Nucleotide Regulatory Protein, mRNA sequence.		
DEFINITION	AV255052		
ACCESSION	AV255052.1 GI:6242511		
VERSION	EST.		
KEYWORDS	Mus musculus (house mouse)		
SOURCE	Mus musculus		
REFERENCE	Unpublished (1999)		
TITLE	Contact: Yoshihide Hayaishizaki		
JOURNAL	Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute		
COMMENT	The Institute of Physical and Chemical Research (RIKEN) 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan Tel: 81-45-503-9222 Fax: 81-45-503-9216 Email: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/ Sasaki,N., Izawa,M., Watahiki,M., Ozawa,K., Tanaka,T., Yoneda,Y., Matsura,S., Carninci,P., Muramatsu,M., Okazaki,Y. and Hayaishizaki,Y. Transcriptional sequencing: A method for DNA sequencing using RNA polymerase. Proc. Natl. Acad. Sci. U.S.A. 95 (7), 3455-3460 (1998) Itou,M., Kitsumi,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J., Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M., Okazaki,Y. and Hayaishizaki,Y. Automated filtration-based high-throughput plasmid preparation system. Genome Res. 9 (5), 463-470 (1999) Carninci,P. and Hayaishizaki,Y. High-efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999) Please visit our web site (http://genome.rtc.riken.go.jp) for further details. Location/Qualifiers 1. .236 /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6J" /db_xref="taxon:10090" /clone="4921516B04" /sex="male" /tissue_type="testis" /dev_stage="adult" /lab_host="DH10B" /clone_lib="RIKEN full-length enriched, adult male testis (DH10B)"		
	/note="Site 1: SalI; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGAGAGAGAGATCCAGAGCTCTTTTTTTTTTTTTTTT 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCGAGTTAAATTAATTCCTCCCTCCCCCCCC 3']. cDNA was cloned into the XhoI and BamHI sites. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI; 3' end: BamHI."		
Query Match	76.8%; Score 19.2; DB 1; Length 236;		
Best Local Similarity 87.5%; Pred. No. 5.4e+02;			
ORIGIN			
Query Match	76.8%; Score 19.2; DB 1; Length 236;		
Best Local Similarity 87.5%; Pred. No. 5.4e+02;			

Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAAATCCTCTAGGCTTTTCAGT 24
||||| ||||| ||||| ||||| |||||
Db 114 AAAAAAATCCTCTAGGCTTTTCATT 91

RESULT 22

BY372216/c
LOCUS
DEFINITION
BY372216 RIKEN full-length enriched, submandibular gland CRL-1734
SCA-9 clone 15 cDNA Mus musculus cDNA G430122K07 3', mRNA
sequence.

ACCESSION

BY372216

VERSION

BY372216.1 GI:26601704

KEYWORDS

EST.

SOURCE

Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

Sciurognathi; Murioidea; Muridae; Murinae; Mus.

1 (bases 1 to 239)

Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S.,

Nikaido,I., Oato,N., Saito,R., Suzuki,H., Yamanaka,I.,

Kiyosawa,H., Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A.,

Schonbach,C., Gojobori,T., Baldarelli,R., Hill,D.P., Bult,C.,

Hume,D.A., Quackenbush,J., Schriml,L.M., Kanapin,A., Matsuda,H.,

Batalov,S., Baisel,K.W., Blake,J.A., Bradt,D., Brusic,V.,

Chethia,C., Corbani,L.E., Cousins,S., Dalla,E., Dragani,T.A.,

Flécher,C.F., Forrest,A., Frazer,K.S., Gaasterland,T.,

Gariboldi,M., Gissi,C., Godzik,A., Gough,J., Grimmond,S.,

Gustincich,S., Hirokawa,N., Jackson,I.J., Jarvis,E.D., Kanai,A.,

Kawaji,H., Kawasawa,Y., Kedzierski,R.M., King,B.L., Konagaya,A.,

Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R.,

Maitais,L., Marchionni,L., McKenzie,L., Miki,H., Nagashima,T.,

Numata,K., Okido,T., Pavan,W.J., Pertea,G., Pesole,G.,

Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramachandran,S.,

Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring,B.Z., Ringwald,M.,

Sandelin,A., Schneider,C., Semple,C.A., Setou,M., Shimada,K.,

Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tomita,M.,

Verardo,R., Wagner,L., Wahlestedt,C., Wang,Y., Watanabe,Y.,

Wells,C., Wilning,L.G., Wynshaw-Boris,A., Yanagisawa,M., Yang,I.,

Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A., Carninci,P.,

Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura,K.,

Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Inotani,K., Ishii,Y.,

Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata,K.,

Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S.,

Rogers,J., Birney,E. and Hayashizaki,Y.

Analysis of the mouse transcriptome based on functional annotation

of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)

12466851

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic

Sciences Center(GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-resgsc.riken.jp, URL:http://genome.gsc.riken.jp/

Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S.,

Hirozane,T., Inotani,K., Ishii,Y., Itoh,M., Kawai,J., Konno,H.,

Miyazaki,A., Murata,M., Nakamura,M., Nomura,K., Numazaki,R.,

Ohno,M., Sakai,K., Sakazume,N., Sasaki,D., Sato,K., Shibata,K.,

Shiraki,T., Tagami,M., Waki,K., Watahiki,A., Muramatsu,M. and

Hayashizaki,Y. Direct Submission

Computational Analysis of Full-Length Mouse cDNAs Compared with

Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new

genes. Genome Res. 10 (10), 1617-1630 (2000)

RIKEN integrated sequence analysis (RISA) system--384-format

sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA

encyclopedia: real-time sequence clustering for construction of a

nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

cDNA library was prepared and sequenced in Mouse Genome

Encyclopedia Project of Genome Exploration Research Group in Riken

Genomic Sciences Center and Genome Science Laboratory in RIKEN.

Division of Experimental Animal Research in Riken contributed to

prepare mouse tissues.

Please visit our web site (<http://genome.gsc.riken.go.jp>) for

further details.

FEATURES

source

Location/Qualifiers

1..239

/organism="Mus musculus"

/mol_type="mRNA"

/strain="Swiss Webster"

/db_xref="taxon:10090"

/clone="G430122K07"

/sex="male"

/tissue type="submandibular gland"

/cell_line="CRL-1734 SCA-9 clone 15"

/clone_lib="RIKEN full-length enriched, submandibular

gland CRL-1734 SCA-9 clone 15 cDNA"

gland CRL-1734 SCA-9 clone 15 cDNA"

Query Match 76.8%; Score 19.2; DB 5; Length 239;

Best Local Similarity 87.5%; Pred. No. 5.4e+02;

Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAAATCCTCTAGGCTTTTCAGT 24

||||| ||||| ||||| ||||| |||||

Db 118 AAAAAAATCCTCTAGGCTTTTCATT 95

||||| ||||| ||||| ||||| |||||

RESULT 23

LOCUS

AV304939/c

DEFINITION

AV304939 RIKEN full-length enriched, 8 days embryo Mus musculus

cDNA clone 570524I16 3', similar to AB003502 Mus musculus mRNA for

Guanine Nucleotide Regulatory Protein, mRNA sequence.

AV304939

VERSION

AV304939.1 GI:6337453

KEYWORDS

Mus musculus (house mouse)

SOURCE

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

Sciurognathi; Murioidea; Muridae; Murinae; Mus.

1 (bases 1 to 251)

Konno,H., Aizawa,K., Akahita,S., Akiyama,J., Carninci,P., Endo,T.,

Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N., Hirozane,T., Hori,F.,

Ishii,Y., Ishikawa,I., Itoh,M., Izawa,M., Kadota,K., Kagawa,I.,

Kai,C., Kawai,J., Kikuchi,N., Kojima,Y., Koya,S., Kusakabe,M.,

Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H., Okazaki,Y.,

Owa,C., Ozawa,Y., Saito,H., Sano,M., Sato,K., Shibata,K.,

Shibata,Y., Shigemoto,Y., Shiraki,T., Sogabe,Y., Sugahara,Y.,

Suzuki,H., Suzuki,H., Takahashi,F., Tateo,M., Tomimaga,N.,

Tsunoda,Y., Watahiki,A., Watanabe,S., Yamamura,T., Yasunishi,A.,

Yokota,T., Yoshiki,A., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.

RIKEN Mouse ESTs (Konno,H., et al. 1999)

Unpublished (1999)

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic

Sciences Center(GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-resgsc.riken.jp, URL:http://genome.gsc.riken.jp/

Sasaki,N., Izawa,M., Watahiki,M., Ozawa,K., Tanaka,T., Yoneda,Y.,

Matsuura,S., Carninci,P., Muramatsu,M., Okazaki,Y. and

Hayashizaki,Y.

cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Cloned through the NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert length: 1060 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 286.
Location/Qualifiers
1. .292
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2499843"
/tissue_type="anaplastic oligodendroglioma"
/lab_host="DH10B"
/clone_lib="NCI CGAP Brn25"
/notes="Organ: brain; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site.1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTCAATCTGAAGTGGAGCGCGCATAGTTTTTTTTTTTTTTTTTTT T 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library is normalized, and was constructed by Bento Soares and M.Fatima Bonaldo."

ORIGIN

Query Match 76.8%; Score 19.2; DB 1; Length 292;
Best Local Similarity 87.5%; Pred. No. 5.7e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAACTCTGTAGGCTTTTCAGT 24
||||| ||||||| |||||
Db 125 AAAAAATTCCTGTAGGTTTTCATT 148

RESULT 40
AI801161
LOCUS
DEFINITION
293 bp mRNA linear EST 06-JUL-1999
to79a06.x1 NCI_CGAP_Gas4 Homo sapiens cDNA clone IMAGE:2184514 3',
mRNA sequence.
AI801161
VERSION
KEYWORDS
SOURCE
Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
1 (bases 1 to 293)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
AUTHORS
TITLE
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL
COMMENT
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Cloned through the NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Seq primer: -40UP from Gibco
High quality sequence stop: 265.
Location/Qualifiers
1. .293
/organism="Homo sapiens"
/mol_type="mRNA"

FEATURES

source
1. .293
/organism="Homo sapiens"
/mol_type="mRNA"

/db_xref="taxon:9606"
/clone="IMAGE:2184514"
/tissue_type="poorly differentiated adenocarcinoma with signet ring cell features"
/lab_host="DH10B"
/clone_lib="NCI_CGAP Gas4"
/note="Organ: stomach; Vector: pCMV-SPORT6; Site.1: SalI; Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.69 kb. Life Technologies catalog #: 11549-011"

ORIGIN

Query Match 76.8%; Score 19.2; DB 1; Length 293;
Best Local Similarity 87.5%; Pred. No. 5.7e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAACTCTGTAGGCTTTTCAGT 24
||||| ||||||| |||||
Db 125 AAAAAATTCCTGTAGGTTTTCATT 148

RESULT 41

AW273587
LOCUS
DEFINITION
294 bp mRNA linear EST 03-JAN-2000
xv67f07.x1 NCI_CGAP_Lu28 Homo sapiens cDNA clone IMAGE:2818213 3',
mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
Homo sapiens (human)
ORGANISM

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
1 (bases 1 to 294)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
AUTHORS
TITLE
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL
COMMENT
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Chris Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: Christa Prange, The I.M.A.G.E. Consortium DNA Sequencing by: Washington University Genome Sequencing Center
Cloned through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Seq primer: -40UP from Gibco
High quality sequence stop: 293.
Location/Qualifiers
1. .294
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2818213"
/tissue_type="two pooled squamous cell carcinomas"
/lab_host="DH10B"
/clone_lib="NCI_CGAP Lu28"
/note="Organ: lung; Vector: pCMV-SPORT6; Site.1: SalI; Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt. Library constructed by Life Technologies."

FEATURES

source
1. .294
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2818213"
/tissue_type="two pooled squamous cell carcinomas"
/lab_host="DH10B"
/clone_lib="NCI_CGAP Lu28"
/note="Organ: lung; Vector: pCMV-SPORT6; Site.1: SalI; Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt. Library constructed by Life Technologies."

ORIGIN

Query Match 76.8%; Score 19.2; DB 1; Length 294;
Best Local Similarity 87.5%; Pred. No. 5.7e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAACTCTGTAGGCTTTTCAGT 24
||||| ||||||| |||||
Db 123 AAAAAATTCCTGTAGGTTTTCATT 146

to Lona-linker LL-Sal4, purified by phenol/chloroform, and separated from free linkers by Centricon 100. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer Sal4-S. The products were purified by phenol/chloroform and Centricon 100. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pSPORT1 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by the standard chemical method. The average insert size is about 2.5 kb. The library was constructed by Yulan Piao (NIA)."

ORIGIN

Query Match 76.8%; Score 19.2; DB 3; Length 298;
Best Local Similarity 87.5%; Pred. No. 5.7e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAAATCTCTAGGCTTTTCAGT 24
||||| ||||||| ||||||| |||||||
Db 123 AAAAAAATCTCTAGGCTTTTCATT 146

RESULT 44

AA450256 300 bp mRNA linear EST 04-JUN-1997
LOCUS z42g02.s1 Soares total_fetus_Nb2HF8_9w Homo sapiens cDNA clone
DEFINITION IMAGE:789170 3', mRNA sequence.

ACCESSION AA450256
VERSION AA450256.1 GI:2164006
KEYWORDS EST.
SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.

REFERENCE

AHLLier L., Allen M., Bowles L., Dubuque T., Geisel G., Jost S.,
Kucaba T., Lacy M., Le N., Lennon G., Marra M., Martin J.,
Moore B., Schellenberg K., Streptoe M., Tan F., Theising B.,
White Y., Wyllie T., Waterston R. and Wilson R.
WashU-Merck EST Project 1997
Unpublished (1997)

TITLE

WashU-Merck EST Project 1997

JOURNAL

Contact: Wilson RK

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available royalty-free through LLNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

Seq primer: -41m13 fwd. ET from Amersham

High quality sequence stop: 291.

FEATURES

Location/Qualifiers

1..300

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="GDB:5986787"

/db_xref="taxon:9606"

/clone="IMAGE:789170"

/dev_stage="8-9 weeks"

/lab_host="DH10B"

/clone_lib="Soares total_fetus_Nb2HF8_9w"

/note="Vector: p7T73D-Fac (Pharmacia) with a modified

polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA

was prepared from mRNA obtained from pooled 8-9 week

(total) fetus material with a Not I - oligo(dT) primer [5'

TGTTACCAACTGAGTGGGCGGCTTAATTTTTTTTTTTT 3'].

Double-stranded cDNA was ligated to Eco RI adaptors

(Pharmacia), digested with Not I and cloned into the Not I

and Eco RI sites of the modified p7T73 vector. Library

went through one round of normalization, and was

constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 76.8%; Score 19.2; DB 1; Length 300;
Best Local Similarity 87.5%; Pred. No. 5.7e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAAATCTCTAGGCTTTTCAGT 24
||||| ||||||| ||||||| |||||||
Db 123 AAAAAAATCTCTAGGCTTTTCATT 146

RESULT 45

BM248034 300 bp mRNA linear EST 07-JUN-2003
LOCUS K0810F03-3 NIA Mouse 8.5-dpc Whole Embryo cDNA Library (Long) Mus
DEFINITION musculus cDNA clone NIA:K0810F03 IMAGE:30079358 3', mRNA sequence.

ACCESSION BM248034
VERSION BM248034.2 GI:31502741
KEYWORDS EST.

SOURCE

Mus musculus (house mouse)

ORGANISM

Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

Piao Y., Ko N.T., Lim M.K. and Ko M.S.H.
Construction of long-transcript enriched cDNA libraries from
submicrogram amounts of total RNAs by a universal PCR amplification
method

Genome Res. 11 (9), 1553-1558 (2001)

11544199

On Dec 17, 2001 this sequence version replaced gi:17883304.

Contact: Dawood B. Dudekula

Laboratory of Genetics

National Institute on Aging/National Institutes of Health

333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA

Email: cdna@lgeun.grc.nia.nih.gov

Plate: K0810 row: F column: 03

Seq primer: -21M13 Forward

High quality sequence stop: 300

POLYA=yes.

Location/Qualifiers

1..300

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="niaEST:K0810F03-3"

/db_xref="taxon:10090"

/clone="NIA:K0810F03 IMAGE:30079358"

/tissue_type="whole embryo including extraembryonic

tissues at 8.5-days postcoitum"

/dev_stage="8.5-days postcoitum"

/lab_host="DH10B"

/clone_lib="NIA Mouse 8.5-dpc Whole Embryo cDNA Library

(Long)"

/note="Vector: pSPORT1 (Invitrogen); Site 1: SalI; Site 2:

NotI; Mouse cDNA project by the Laboratory of Genetics,

National Institute on Aging (NIA), Intramural Research

Program, NIH (<http://lgeun.grc.nia.nih.gov/cDNA>). This is

a long-transcript enriched cDNA library (Ref. Genome Res.

11: 1553-1558 (2001). [PMID: 11544199]]. Total RNAs were

extracted from a pool of 13 embryos at 8.5-days

postcoitum. Double-stranded cDNAs were synthesized with an

Oligo(dT) primer [Invitrogen:

5'-PGACTAGTCTAGATCGAGCGCGCTTTT-3'] from

9.1 ug of total RNA, treated with T4 DNA polymerase, and

purified by ethanol-precipitation. The cDNAs were ligated

to Lona-linker LL-Sal4, purified by phenol/chloroform, and

separated from free linkers by Centricon 100. Then, the

cDNAs were amplified by long-range high fidelity PCR using

Ex Taq polymerase (Takara) with a primer Sal4-S. The

products were purified by phenol/chloroform and Centricon

100. The cDNAs were digested with SalI and NotI enzymes

and cloned into SalI/NotI site of pSPORT1 plasmid vector.

The DH10B E. coli host was transformed with the ligation mixture by the standard chemical method. The average insert size is about 2.5 kb. The library was constructed by Yulan Piao (NIA)."

ORIGIN

Query Match 76.8%; Score 19.2; DB 3; Length 300;
Best Local Similarity 87.5%; Pred. No. 5.7e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAAATCCTGTAGGCTTTTCAGT 24
||||| ||||||| ||||||| |||||||
Db 123 AAAAAAATCCTGTAGGCTTTTCATT 146

RESULT 46

AV121807/c
LOCUS
DEFINITION AV121807 Mus musculus C57BL/6J 10-day embryo Mus musculus cdna EST 01-JUL-1999
clone 2610509N22, mRNA sequence.

ACCESSION AV121807

VERSION

KEYWORDS

SOURCE AV121807.1 GI:5307521

ORGANISM

Mus musculus (house mouse)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.

REFERENCE

AUTHORS

1 (bases 1 to 304)
Carninci, P., Shibata, K., Ozawa, Y., Konno, H., Itoh, M., Aizawa, K.,
Akahira, S., Akiyama, J., Fukuda, S., Fukunishi, Y., Funayama, T.,
Hara, A., Hayatsu, N., Hori, F., Ishikawa, T., Itoh, M., Izawa, M.,
Kawai, J., Kikuchi, N., Kojima, Y., Matsuyama, T., Nitsume, H., Oda, H.,
Owa, C., Sato, K., Shibata, Y., Shigemoto, Y., Shiraki, T., Sogabe, Y.,
Sugahara, Y., Suzuki, H., Suzuki, H., Tateno, M., Tomaru, Y.,
Tomimaga, N., Watanabe, S., Yagame, M., Yamamura, T., Yokota, T.,
Yoshino, M., Muramatsu, M., Okazaki, Y. and Hayashizaki, Y.

TITLE

JOURNAL

COMMENT

Unpublished (1999)

Contact: Chie Owa

Genome Science Laboratory

RIKEN

3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan

Tel: 81-298-36-9145

Fax: 81-298-36-9098

Email: genome-resortc.riken.go.jp

Thermostabilization and thermoactivation of thermolabile enzymes by
trehalose and its application for the synthesis of full length cdna
(Proc. Natl. Acad. Sci. U.S.A. 95(2):520-524 (1998))
Transcriptional sequencing: A method for DNA sequencing using RNA
polymerase (Proc. Natl. Acad. Sci. U.S.A. 95(7):3455-3460 (1998))
Please visit our web site (http://genome.rtc.riken.go.jp) for
further details.

FEATURES

source

1..304
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="2610509N22"
/sex="mixed"
/dev_stage="10-day embryo"
/clone_lib="Mus musculus C57BL/6J 10-day embryo"

ORIGIN

Query Match 76.8%; Score 19.2; DB 1; Length 304;
Best Local Similarity 87.5%; Pred. No. 5.7e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAAATCCTGTAGGCTTTTCAGT 24

||||| ||||||| ||||||| |||||||

Db 170 AAAAAAATCCTGTAGGCTTTTCATT 147

RESULT 47

LOCUS

DEFINITION

AI869356
tw39b11.x1 NCI_CGAP_Utl1 Homo sapiens cdna clone IMAGE:2262045 3',
mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

AI869356
AI869356.1 GI:5543324
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1 (bases 1 to 310)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.

CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone Distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html

Insert Length: 1767 Std Error: 0.00

Seq primer: -40UP from Gibco

High quality sequence stop: 294.

Location/Qualifiers

1..310
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2262045"
/tissue_type="well-differentiated endometrial
adenocarcinoma, 7 pooled tumors"
/lab_host="DH10B"
/clone_lib="NCI CGAP_Utl1"
/note="Organ: uterus; Vector: pCMV-SPORT6; Site.1: SalI;
Site.2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.75 kb. Life Technologies catalog #:
11538-014"

ORIGIN

Query Match 76.8%; Score 19.2; DB 1; Length 310;
Best Local Similarity 87.5%; Pred. No. 5.7e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAAATCCTGTAGGCTTTTCAGT 24

||||| ||||||| ||||||| |||||||

Db 136 AAAAAAATCCTGTAGGCTTTTCATT 159

RESULT 48

LOCUS

DEFINITION

AA117935
mn08a04.r1 Beddington mouse embryonic region Mus musculus cdna
clone IMAGE:537294 5', mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

AA117935
AA117935.1 GI:1672959
EST.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.

REFERENCE

AUTHORS

1 (bases 1 to 311)
Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,

Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R.

TITLE

JOURNAL COMMENT

The WashU-HMI Mouse EST Project
 Contact: Marra M/Mouse EST Project
 WashU-HMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu

This clone is available royalty-free through LLMNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:324230

Putative full length read

vector to vector length is 390

Seq primer: -40ml3 ET

High quality sequence stop: 310.

FEATURES

source

1. .311

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6 x DBA"

/db_xref="taxon:10090"

/clone="IMAGE:537294"

/sex="pooled"

/tissue_type="embryo"

/dev_stage="7.5dpc"

/lab_host="DH12S"

/clone_lib="Beddington mouse embryonic region"

/note="Organ: whole embryo; Vector: pCMV-SPORT; Site: 1:

Sal; Site 2: NotI; Cloned unidirectionally. Primer:

Oligo dT. Gastrulating embryos were collected at 7.5dpc

from C57BL/6 x DBA matings, excluding embryos that had

developed head folds and all extraembryonic tissues.

Average insert size: 1.3 kb (range: 0.5 - 3.0 kb).

Referenced in Development 121, 2479-2489 (1995)"

ORIGIN

Query Match

Best Local Similarity 76.8%; Score 19.2; DB 1; Length 311;

Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AAAAACTCTGTAGGCTTTCAGT 24

||||| ||||| ||||| ||||| |||||

Db 267 AAAAAATCTGTAGGTTTTCATT 244

RESULT 49

AW323152

LOCUS

DEFINITION u058c11.x1 NCI_CGAP_Lu29 Mus musculus cDNA clone IMAGE:2646740 3', mRNA sequence.

ACCESSION AW323152

VERSION AW323152.1 GI:6757177

KEYWORDS EST.

SOURCE Mus musculus

ORGANISM Mus musculus (house mouse)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 312)

NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

AUTHORS Tumor Gene Index

TITLE Unpublished (1997)

JOURNAL Contact: Robert Strausberg, Ph.D.

COMMENT Email: cgapbs-re@mail.nih.gov

Tissue Procurement: Gilbert Smith, Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html

MGI:1027192

Seq primer: -40UP from Gibco

High quality sequence stop: 311.

FEATURES

source

Location/Qualifiers

1. .312

/organism="Mus musculus"

/mol_type="mRNA"

/strain="Czech II"

/db_xref="taxon:10090"

/clone="IMAGE:2646740"

/tissue_type="spontaneous tumor, metastatic to mammary.

Stem cell origin."

/lab_host="DH108"

/clone_lib="NCI_CGAP Lu29"

/note="Organ: lung; Vector: pCMV-SPORT6; Site_1: SalI;

Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.

Library constructed by Life Technologies. Investigator

providing samples: Gilbert Smith, NIH"

ORIGIN

Query Match

Best Local Similarity 76.8%; Score 19.2; DB 1; Length 312;

Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AAAAACTCTGTAGGCTTTCAGT 24

||||| ||||| ||||| ||||| |||||

Db 126 AAAAAATCTGTAGGTTTTCATT 149

RESULT 50

CX846858

LOCUS

DEFINITION

CX846858 313 bp mRNA linear EST 02-FEB-2005

JGI_CAAK12911.rev NIH XGC tropBn3 Xenopus tropicalis cDNA clone

IMAGE:7793462 3', mRNA sequence.

ACCESSION CX846858

VERSION CX846858.1 GI:58502125

KEYWORDS EST.

SOURCE Xenopus tropicalis (western clawed frog)

ORGANISM Xenopus tropicalis

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;

Xenopodinae; Xenopus; Silurana.

1 (bases 1 to 313)

Richardson, P., Lucas, S., Rokhsar, D., Dettter, J.C., NG, D.C.,

Brokstein, P. and Lindquist, E.A.

DOE Joint Genome Institute Xenopus tropicalis EST project

Unpublished (2004)

Other_ESTs: JGI_CAAK12911.fwd

Contact: Lindquist, E.A., Richardson, P.

DOE Joint Genome Institute

2800 Mitchell Drive, Walnut Creek, CA 94598, USA

Tel: 925 296 5600

Fax: 925 296 5710

Email: cdna@jgi-psf.org

Tissue Procurement: Timothy Grammer (Richard M. Harland Laboratory,

University of California, Berkeley:

<http://tropicalis.berkeley.edu/home>

cDNA Library Preparation: DOE Joint Genome Institute:

<http://www.jgi.doe.gov>

DNA Sequencing: DOE Joint Genome Institute: <http://www.jgi.doe.gov>

Clone Distribution: I.M.A.G.E. Consortium/LLNL:

<http://image.llnl.gov>

Naming Conventions: EST name is generated by the concatenation of

the JGI Clone id and the direction of sequencing. The suffix 'rev'

indicates a reverse sequencing read of the insert. It does not

necessarily reflect the orientation of the insert.

Poly-A: Based upon the presence of a run of 14 or more T residues

at the beginning of the sequence, this clone was polyadenylated.

The resulting Poly-T sequence has been removed.

Plate: CAAK 0133 row: m column: 12
High quality sequence stop: 311
POLYA=Yes.

FEATURES
 source
 Location/Qualifiers
 1..313
 /organism="Xenopus tropicalis"
 /mol_type="mRNA"
 /db_xref="taxon:8364"
 /clone="IMAGE:7793462"
 /tissue_type="Brain"
 /dev_stage="Adult"
 /lab_host="ElectroMAX DH10B"
 /clone_lib="NIH_XGC_tropBm3"
 /note="Vector: pCMVSPORT6; Site_1: SalI; Site_2: NotI;
 This library was made from dr primed cDNA and cloned into
 Invitrogen pCMVSPORT6 vector. The work was done at DOE
 Joint Genome Institute. Poly A RNA were primed with 5',
 GACTAGTCTAGATCGCAG CGGCGCCCTTTT TTTT 3'. cDNA
 were ligated to SalI adapter (5' TCGACCCACGCGTCCG and
 5'CGACGCGTGGG), digested with NotI, size fractionated in
 1.1% agarose gel electrophoresis and ligated into NotI and
 SalI digested pCMVSPORT6 vector."

ORIGIN

Query Match 76.8%; Score 19.2; DB 8; Length 313;
Best Local Similarity 87.5%; Pred. No. 5.7e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAACTCCTGTAGGCTTTCAGT 24
 ||||| ||||| ||||| |||||
Db 252 AAAAACTCCTGTAGGCTTTCAGT 275

Search completed: February 3, 2006, 22:02:21
Job time : 2960.67 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 3, 2006, 11:35:31 ; Search time 80.6667 Seconds
(without alignments)
550.897 Million cell updates/sec

Title: US-10-719-900-20
Perfect score: 25
Sequence: 1 aaaaactctgtaggcttcagt 25

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues 2606114

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 150 summaries

Database :

1: /cgn2_6/ptodata/1/ina/1 COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5 COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B COMB.seq:*
5: /cgn2_6/ptodata/1/ina/H COMB.seq:*
6: /cgn2_6/ptodata/1/ina/PTUS COMB.seq:*
7: /cgn2_6/ptodata/1/ina/PP COMB.seq:*
8: /cgn2_6/ptodata/1/ina/RE COMB.seq:*
9: /cgn2_6/ptodata/1/ina/backfileseq1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	19.2	76.8	2584	3	US-09-949-016-5287
C 2	19.2	76.8	47284	3	US-09-949-016-17029
C 3	18.8	75.2	23776	3	US-09-949-002-702
4	18.6	74.4	601	3	US-09-949-016-76001
5	18.6	74.4	601	3	US-09-949-016-76002
6	18.6	74.4	39982	3	US-09-820-924-3
7	18.6	74.4	39982	3	US-10-369-626-3
8	18.6	74.4	51043	3	US-09-948-016-12739
9	18.6	74.4	51046	3	US-09-949-016-13946
10	17.6	70.4	1563	2	US-08-279-590A-1
11	17.6	70.4	1563	3	US-08-910-092-1
12	17.6	70.4	1657	3	US-09-270-767-9840
13	17.6	70.4	1867	3	US-08-630-915A-13
14	17.6	70.4	1867	3	US-09-879-357-13
15	17.6	70.4	8511	3	US-09-949-016-14976
16	17.6	70.4	36816	3	US-09-949-016-16218
C 17	17.6	70.4	118143	3	US-09-949-016-17196
C 18	17.2	68.8	601	3	US-09-949-016-144942
19	17.2	68.8	601	3	US-09-949-016-144943
C 20	17.2	68.8	1971	3	US-09-620-312D-960
C 21	17.2	68.8	3914	3	US-10-104-047-850
C 22	17.2	68.8	52667	3	US-09-949-016-12019
C 23	17.2	68.8	146307	3	US-09-949-016-14881
24	17.2	68.8	146307	3	US-09-949-016-14882

25	17.2	68.8	146307	3	US-09-949-016-14883	Sequence 14883, A
26	17.2	68.8	146307	3	US-09-949-016-14884	Sequence 14884, A
27	17.2	68.8	146307	3	US-09-949-016-14885	Sequence 14885, A
28	17.2	68.8	146307	3	US-09-949-016-14886	Sequence 14886, A
29	17.2	68.8	146307	3	US-09-949-016-14887	Sequence 14887, A
30	17.2	68.8	146307	3	US-09-949-016-14888	Sequence 14888, A
31	17.2	68.8	148405	3	US-09-949-016-11747	Sequence 11747, A
32	17.2	68.8	148405	3	US-09-949-016-12835	Sequence 12835, A
33	17.2	68.8	148405	3	US-09-949-016-12836	Sequence 12836, A
34	17.2	68.8	148405	3	US-09-949-016-12837	Sequence 12837, A
C 35	17.2	68.8	169998	3	US-09-676-610B-24	Sequence 24, Appl
C 36	17.2	68.8	177251	3	US-09-949-016-15841	Sequence 15841, A
C 37	17.2	68.8	197496	3	US-09-877-177A-10	Sequence 10, Appl
C 38	17	68.0	601	3	US-09-949-016-133546	Sequence 133546, A
C 39	17	68.0	601	3	US-09-949-016-133547	Sequence 133547, A
40	17	68.0	703	3	US-09-270-767-982	Sequence 982, App
41	17	68.0	703	3	US-09-270-767-16264	Sequence 16264, A
C 42	17	68.0	1095	3	US-09-134-000C-625	Sequence 625, App
C 43	17	68.0	1939	3	US-09-799-451-566	Sequence 566, App
C 44	17	68.0	2905	3	US-09-949-016-1046	Sequence 1046, App
C 45	17	68.0	3416	3	US-09-949-016-4827	Sequence 4827, App
C 46	17	68.0	3416	3	US-09-949-016-4828	Sequence 4828, App
C 47	17	68.0	3490	3	US-09-949-016-4829	Sequence 4829, App
C 48	17	68.0	3490	3	US-09-949-016-4830	Sequence 4830, App
49	17	68.0	17132	3	US-09-949-016-15361	Sequence 15361, A
50	17	68.0	24150	3	US-09-949-016-12438	Sequence 12438, A
C 51	17	68.0	30271	3	US-09-949-016-12796	Sequence 12796, A
C 52	17	68.0	30272	3	US-09-949-016-14006	Sequence 14006, A
C 53	17	68.0	36075	3	US-09-949-016-16571	Sequence 16571, A
C 54	17	68.0	36075	3	US-09-949-016-16572	Sequence 16572, A
C 55	17	68.0	36625	3	US-09-949-016-12788	Sequence 12788, A
C 56	17	68.0	37133	3	US-09-949-016-16569	Sequence 16569, A
C 57	17	68.0	37133	3	US-09-949-016-16570	Sequence 16570, A
C 58	17	68.0	202111	3	US-09-949-016-13877	Sequence 13877, A
C 59	17	68.0	245286	3	US-09-949-016-15497	Sequence 15497, A
60	17	68.0	256171	3	US-09-949-016-12822	Sequence 12822, A
61	17	68.0	256176	3	US-09-949-016-15224	Sequence 15224, A
62	17	68.0	451924	3	US-09-949-016-12896	Sequence 12896, A
63	17	68.0	451925	3	US-09-949-016-17305	Sequence 17305, A
64	16.8	67.2	442	3	US-09-270-767-12773	Sequence 12773, A
C 65	16.8	67.2	1589	3	US-09-799-451-637	Sequence 637, App
C 66	16.8	67.2	8341	2	US-08-737-825-1	Sequence 1, Appl
C 67	16.6	66.4	306	3	US-09-513-999C-8250	Sequence 8250, App
68	16.6	66.4	316	3	US-09-662-402A-28	Sequence 28, Appl
69	16.6	66.4	601	3	US-09-949-016-31372	Sequence 31372, A
70	16.6	66.4	601	3	US-09-949-016-76138	Sequence 76138, A
71	16.6	66.4	601	3	US-09-949-016-122335	Sequence 122335, A
72	16.6	66.4	601	3	US-09-949-016-135598	Sequence 135598, A
C 73	16.6	66.4	601	3	US-09-949-016-148083	Sequence 148083, A
C 74	16.6	66.4	601	3	US-09-949-016-148189	Sequence 148189, A
C 75	16.6	66.4	601	3	US-09-949-016-204052	Sequence 204052, A
C 76	16.6	66.4	2660	3	US-09-949-016-2206	Sequence 2206, App
77	16.6	66.4	2669	3	US-09-949-016-691	Sequence 691, App
78	16.6	66.4	2674	2	US-08-926-724-2	Sequence 2, Appl
C 79	16.6	66.4	4530	3	US-09-221-017B-913	Sequence 913, App
C 80	16.6	66.4	4756	3	US-09-949-016-15185	Sequence 15185, A
C 81	16.6	66.4	19299	3	US-09-949-016-13014	Sequence 13014, A
82	16.6	66.4	52523	3	US-09-949-016-12433	Sequence 12433, A
83	16.6	66.4	52530	3	US-09-949-016-13948	Sequence 13948, A
C 84	16.6	66.4	98701	3	US-09-949-016-15998	Sequence 15998, A
C 85	16.6	66.4	98701	3	US-09-949-016-15999	Sequence 15999, A
C 86	16.6	66.4	129380	3	US-09-949-016-12544	Sequence 12544, A
C 87	16.6	66.4	139257	3	US-09-920-671-11	Sequence 11, Appl
C 88	16.6	66.4	161124	3	US-09-949-016-11760	Sequence 11760, A
C 89	16.6	66.4	187595	3	US-09-949-016-15846	Sequence 15846, A
90	16.6	66.4	276237	3	US-09-949-016-17504	Sequence 17504, A
91	16.6	66.4	392000	3	US-10-027-983-11	Sequence 11, Appl
C 92	16.4	65.6	269	3	US-09-513-999C-9119	Sequence 9119, App
C 93	16.4	65.6	601	3	US-09-949-016-149843	Sequence 149843, A
94	16.4	65.6	89450	3	US-09-949-016-15848	Sequence 15848, A
C 95	16.4	65.6	105050	3	US-09-949-016-15953	Sequence 15953, A
C 96	16.2	64.8	450	3	US-09-513-999C-28047	Sequence 28047, A
C 97	16.2	64.8	601	3	US-09-949-016-32668	Sequence 32668, A

US-09-949-002-702

```
Query Match      75.2%; Score 18.8; DB 3; Length 23776;
Best Local Similarity 90.9%; Pred. No. 33;
Matches 20: Conservative 0; Mismatches 2; Indels 0;
```

Qy 4 A A A C T C C T G T A G G C T T T C A G T G 25
1282 A A A A T C C T G G A G G C T T T C A G T G 1261

RESULT 4

US-09-949-016-76001
; Sequence 76001, Application US/09949016
; Patent No. 6812339

; GENERAL INFORMATION:

```

, APPLICANT: VENTER, J. Craig et al.
, TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
, TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
, FILE REFERENCE: CL001307
, CURRENT APPLICATION NUMBER: US/09/949,016
, CURRENT FILING DATE: 2000-04-14
, PRIOR APPLICATION NUMBER: 60/241,755
, PRIOR FILING DATE: 2000-10-20
, PRIOR APPLICATION NUMBER: 60/237,768
, PRIOR FILING DATE: 2000-10-03
, PRIOR APPLICATION NUMBER: 60/231,498
, PRIOR FILING DATE: 2000-09-08
, NUMBER OF SEQ ID NOS: 207012
, SOFTWARE: FastSeq for Windows Version 4.0
, SEQ ID NO 76001
, LENGTH: 601

```

```

; LENGTH: 601
TIME DAY

```

; TYPE: DNA
 ; ORGANISM: Human
 US-09-949-016-76001

Query Match 74.4%; Score 18.6; DB 3; Length 601;
Best Local Similarity 84.0%; Pred. No. 16;
Matches 21: Conservative 0; Mismatches 4; Indels

Qy

1 AAAAACTCCTGTAGGCTTTCAGTG 25
||| ||| ||| ||| ||| ||| ||| |||

pB

442 AAAGAAATCCTGTGGCTTTCGTG 466
||||| ||| ||| ||| ||| ||| ||| |||

RESIST 5

US-09-949-016-76002
; Sequence 76002, Application US/09949016

; Patent No. 6812339

```

; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 76002
; LENGTH: 601

```

: LENGTH: 601

TYPE: DNA
ORGANISM: Human

Query Match 74.4%; Score 18.6; DB 3; Length 601;
Best Local Similarity 84.0%; Pred. No. 16;
Matches 21: Conservative 0; Mismatches 4; Indels

QY
1 AAAAAACTCCTGTAGGCTTTTCAGTG 25
pb
61 AAAGAAATCCTGTTGGCTTTCTGTG 85

RESULT 6

US-09-820-924-3
; Sequence 3, Application US/09820924
; Patent No. 6555351

GENERAL INFORMATION:

```

1  APPLICANT: BEASLEY, Ellen M. et al
2  TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
3  TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
4  TITLE OF INVENTION: THEREOF
5  FILE REFERENCE: CL001213
6  CURRENT APPLICATION NUMBER: US/09/820,924
7  CURRENT FILING DATE: 2001-03-30
8  NUMBER OF SEQ ID NOS: 4
9  SOFTWARE: Fast-SEQ for Windows Version 4.0
10 SEQ ID NO 3
11 LENGTH: 39982
12 TYPE: DNA
13 ORGANISM: Human
14 US-09-820-924-3

```

Query Match

Best Local Similarity 84.0%: Pred. No. 48:

Mismatches	0;	Mismatches	4;	Indels	0;	Gaps	0;
Matches	21:	Conservative					

QY 1 AAAAAACTCCTGTAGGCTTTCAGTG 25
||| || ||||| ||||| ||||| |||||
Db 14081 AAAGAAATCCTGTTGGCTTTCGTG 14

RESULT, T 7

US-10-369-626-3
; Sequence 3, Application US/10369626
; Patent No. 6787344

GENERAL INFORMATION:

```

; APPLICANT: BEASLEY, Ellen M. et al
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001213DIV
; CURRENT APPLICATION NUMBER: US/10/369,626
; CURRENT FILING DATE: 2003-02-21
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 39982
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-369-626-3

```

Query Match	74.4%	Score 18.6	DB 3:	Length 39982
-------------	-------	------------	-------	--------------

Query matrix
Best local similarity 84.0%: Pred. No. 48;

Matches	21	Conservative	0	Mismatches	4	Indels	0	Gaps	0
---------	----	--------------	---	------------	---	--------	---	------	---

Qy 1 AAAAACTCCTGTAGGCTTTCAGTG 25
||| ||| ||| ||| ||| ||| ||| |||
nb 14081 AAAGAAATCCTGTGGCTTTCGTG 14

RESIST. T. 8

US-09-949-016-12739
; Sequence 12739, Application US/09949016

; Patent No. 6812339

```

; GENERAL INFORMATION:
;
; APPLICANT: VENTER, J. Craig et al.
;
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;
; FILE REFERENCE: CL001307

```

```
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12739
; LENGTH: 51043
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(51043)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12739
```

```
Query Match 74.4%; Score 18.6; DB 3; Length 51043;
Best Local Similarity 84.0%; Pred. No. 51;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
```

```
QY 1 AAAAACTCCTGTAGGCTTTTCAGTG 25
Db 13899 AAAGAAATCCTGTGGCTTTTCGTG 13923
```

RESULT 9

```
US-09-949-016-13946
; Sequence 13946, Application US/09949016
; Patent No. 6812339
```

GENERAL INFORMATION:

```
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13946
; LENGTH: 51046
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
```

```
; NAME/KEY: misc_feature
; LOCATION: (1)...(51046)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13946
```

```
Query Match 74.4%; Score 18.6; DB 3; Length 51046;
Best Local Similarity 84.0%; Pred. No. 51;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
```

```
QY 1 AAAAACTCCTGTAGGCTTTTCAGTG 25
Db 13899 AAAGAAATCCTGTGGCTTTTCGTG 13923
```

RESULT 10

```
US-08-279-590A-1
; Sequence 1, Application US/08279590A
; Patent No. 5656749
; GENERAL INFORMATION:
; APPLICANT: ELIOT R. SPINDEL, SRINIVASA
```

```
; APPLICANT: NAGALLA AND BRENDA BARRY
; TITLE OF INVENTION: NUCLEIC ACIDS ENCODING
; RECEPTORS FOR BOMBESIN-LIKE
; PEPTIDES
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 50Z or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/279,590A
; FILING DATE: July 22, 1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: No. 5656749e
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Y. Rocky Tsao
; REGISTRATION NUMBER: 34,053
; REFERENCE/DOCKET NUMBER: 00537/098001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1563
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-279-590A-1
```

```
Query Match 70.4%; Score 17.6; DB 2; Length 1563;
Best Local Similarity 83.3%; Pred. No. 66;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
```

```
QY 1 AAAAACTCCTGTAGGCTTTTCAGT 24
Db 187 AAAAAATCGCTGAGGTTTTCAGT 210
```

RESULT 11

```
US-08-910-092-1
; Sequence 1, Application US/08910092
; Patent No. 5814463
```

GENERAL INFORMATION:

```
; APPLICANT: Spindel, Eliot R.
; APPLICANT: Nagalla, Srinivasa
; APPLICANT: Barry, Brenda
; TITLE OF INVENTION: NUCLEIC ACIDS ENCODING RECEPTORS
; FOR BOMBESIN-LIKE PEPTIDES
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
```

APPLICATION NUMBER: US/08/910,092
FILING DATE: 12-AUG-1997
PRIOR APPLICATION DATA: 08/279,590
FILING DATE: 22-JUL-1994
ATTORNEY/AGENT INFORMATION:
NAME: Tsao, Y. Rocky
REGISTRATION NUMBER: 34,053
REFERENCE/DOCKET NUMBER: 00537/098002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1563 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 192...1319
US-08-910-092-1

Query Match 70.4%; Score 17.6; DB 2; Length 1563;
Best Local Similarity 83.3%; Pred. No. 66;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAACTCTGTAGGCTTTTCAGT 24
||||| ||||| ||||| ||||| |||||
DB 187 AAAAAATGCTGAAGGTTTTCAGT 210

RESULT 12

US-09-270-767-9840
Sequence 9840, Application US/09270767
Patent No. 6703491

GENERAL INFORMATION:
APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
FILE REFERENCE: File Reference: 7326-094
CURRENT APPLICATION NUMBER: US/09/270,767
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 62517
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 9840
LENGTH: 1657
TYPE: DNA
ORGANISM: Drosophila melanogaster
US-09-270-767-9840

Query Match 70.4%; Score 17.6; DB 3; Length 1657;
Best Local Similarity 83.3%; Pred. No. 67;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAACTCTGTAGGCTTTTCAGT 24
||||| ||||| ||||| ||||| |||||
DB 550 AGAAAACTCTGTGGGCTATCAGT 573

RESULT 13

US-08-630-915A-13
Sequence 13, Application US/08630915A
Patent No. 6309820

GENERAL INFORMATION:
APPLICANT: SPARKS, Andrew B.
APPLICANT: HOFFMAN, No. 6309820h
APPLICANT: KAY, Brian K.
APPLICANT: FOWLES, Dana M.
TITLE OF INVENTION: POLYPEPTIDES HAVING A FUNCTIONAL
TITLE OF INVENTION: DOMAIN OF INTEREST AND METHODS OF IDENTIFYING AND
...TITLE OF INVENTION: USING SAME

NUMBER OF SEQUENCES: 227
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/630,915A
FILING DATE: 03-APR-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Misrock, S. Lealie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 1101-174
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 1867 bases
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-630-915A-13

Query Match 70.4%; Score 17.6; DB 3; Length 1867;
Best Local Similarity 83.3%; Pred. No. 69;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAACTCTGTAGGCTTTTCAGT 24
||||| ||||| ||||| ||||| |||||
DB 750 ATAAAGTCTGTGGCTTTTCAGT 773

RESULT 14

US-09-879-957-13
Sequence 13, Application US/09879957
Patent No. 6709821

GENERAL INFORMATION:
APPLICANT: SPARKS, Andrew B.
APPLICANT: HOFFMAN, No. 6709821h
APPLICANT: KAY, Brian K.
APPLICANT: FOWLES, Dana M.
APPLICANT: MCCONNELL, Stephen J.

TITLE OF INVENTION: POLYPEPTIDES HAVING A FUNCTIONAL
TITLE OF INVENTION: DOMAIN OF INTEREST AND METHODS OF IDENTIFYING AND
USING SAME

NUMBER OF SEQUENCES: 227
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/879,957
FILING DATE: 13-Jun-2001
CLASSIFICATION: <Unknown>

```
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/630,915
; FILING DATE: 03-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Misrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 1101-174
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1867 bases
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 13:
US-09-879-957-13

Query Match 70.4%; Score 17.6; DB 3; Length 1867;
Best Local Similarity 83.3%; Pred. No. 69;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAACTCTGTAGCTTTTCAGT 24
Db 750 ATAAAGTCTCTGGGCTTTGAGT 773

RESULT 15
US-09-949-016-14976
; Sequence 14976, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14976
; LENGTH: 8511
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-14976

Query Match 70.4%; Score 17.6; DB 3; Length 8511;
Best Local Similarity 83.3%; Pred. No. 1e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAACTCTGTAGCTTTTCAGT 24
Db 5581 AAAAACCTCTGTAATCTTTTCATT 5604

RESULT 16
US-09-949-016-16218/c
; Sequence 16218, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
```

```
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16218
; LENGTH: 36816
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-16218

Query Match 70.4%; Score 17.6; DB 3; Length 36816;
Best Local Similarity 83.3%; Pred. No. 1.5e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAACTCTGTAGGCTTTTCAGT 24
Db 36272 AACATACTCTGTAGTCTTTTCATT 36249

RESULT 17
US-09-949-016-17196/c
; Sequence 17196, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17196
; LENGTH: 118143
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17196

Query Match 70.4%; Score 17.6; DB 3; Length 118143;
Best Local Similarity 83.3%; Pred. No. 2e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAACTCTGTAGGCTTTTCAGT 24
Db 8436 AAAAAATCCTATAGCATTTCAGT 8413

RESULT 18
US-09-949-016-144942
; Sequence 144942, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
```

```
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 144942
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-144942

Query Match      68.8%; Score 17.2; DB 3; Length 601;
Best Local Similarity 86.4%; Pred. No. 82;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAACTCTGTAGGCTTTCA 22
   ||||| ||||| ||||| |||||
DB 369 AAAAACTCTGTAAACATTCA 390

RESULT 19
US-09-949-016-144943
; Sequence 144943, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 144943
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-144943

Query Match      68.8%; Score 17.2; DB 3; Length 601;
Best Local Similarity 86.4%; Pred. No. 82;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAACTCTGTAGGCTTTCA 22
   ||||| ||||| ||||| |||||
DB 2 AAAAACTCTGTAAACATTCA 23

RESULT 20
US-09-620-312D-960/c
; Sequence 960, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunqing
; APPLICANT: Wang, Dunrui
```

```
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 960
; LENGTH: 1971
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1884)
US-09-620-312D-960

Query Match      68.8%; Score 17.2; DB 3; Length 1971;
Best Local Similarity 86.4%; Pred. No. 1.1e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAACTCTGTAGGCTTTCA 22
   ||||| ||||| ||||| |||||
DB 1403 AAAAAACCTGTGAAGCTTTGA 1382

RESULT 21
US-10-104-047-850/c
; Sequence 850, Application US/10104047
; Patent No. 6943241
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. 6943241el full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 850
; LENGTH: 3914
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-104-047-850

Query Match      68.8%; Score 17.2; DB 3; Length 3914;
Best Local Similarity 86.4%; Pred. No. 1.3e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAACTCTGTAGGCTTTCA 22
   ||||| ||||| ||||| |||||
DB 1031 AAACACTCTGTGGCCTTTCA 1010

RESULT 22
US-09-949-016-12019/c
; Sequence 12019, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
```



```
; ORGANISM: Human
US-09-949-016-14884

Query Match          68.8%; Score 17.2; DB 3; Length 146307;
Best Local Similarity 86.4%; Pred. No. 3.3e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAACTCTGTAGGCTTTTCAG 23
Db 18431 AAGACCTCTGTGGCTTTTCAG 18452

RESULT 27
US-09-949-016-14885
; Sequence 14885, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14885
; LENGTH: 146307
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-14885

Query Match          68.8%; Score 17.2; DB 3; Length 146307;
Best Local Similarity 86.4%; Pred. No. 3.3e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAACTCTGTAGGCTTTTCAG 23
Db 18431 AAGACCTCTGTGGCTTTTCAG 18452

RESULT 28
US-09-949-016-14886
; Sequence 14886, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14886
; LENGTH: 146307
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-14886

Query Match          68.8%; Score 17.2; DB 3; Length 146307;
Best Local Similarity 86.4%; Pred. No. 3.3e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAACTCTGTAGGCTTTTCAG 23
Db 18431 AAGACCTCTGTGGCTTTTCAG 18452

RESULT 29
US-09-949-016-14887
; Sequence 14887, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14887
; LENGTH: 146307
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-14887

Query Match          68.8%; Score 17.2; DB 3; Length 146307;
Best Local Similarity 86.4%; Pred. No. 3.3e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAACTCTGTAGGCTTTTCAG 23
Db 18431 AAGACCTCTGTGGCTTTTCAG 18452

RESULT 30
US-09-949-016-14888
; Sequence 14888, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14888
; LENGTH: 146307
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-14888

Query Match          68.8%; Score 17.2; DB 3; Length 146307;
Best Local Similarity 86.4%; Pred. No. 3.3e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAACTCTGTAGGCTTTTCAG 23
Db 18431 AAGACCTCTGTGGCTTTTCAG 18452
```

```
RESULT 31
US-09-949-016-11747
; Sequence 11747, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11747
; LENGTH: 148405
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-11747

Query Match      68.8%; Score 17.2; DB 3; Length 148405;
Best Local Similarity 86.4%; Pred. No. 3.4e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAAGCTCTGTAGGCTTTTCAG 23
Db 18431 AAGACCTCTGTGTGGCTTTTCAG 18452

RESULT 32
US-09-949-016-12835
; Sequence 12835, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12835
; LENGTH: 148405
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12835

Query Match      68.8%; Score 17.2; DB 3; Length 148405;
Best Local Similarity 86.4%; Pred. No. 3.4e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAAGCTCTGTAGGCTTTTCAG 23
Db 18431 AAGACCTCTGTGTGGCTTTTCAG 18452

RESULT 33
US-09-949-016-12836
; Sequence 12836, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12836
; LENGTH: 148405
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12836

Query Match      68.8%; Score 17.2; DB 3; Length 148405;
Best Local Similarity 86.4%; Pred. No. 3.4e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAAGCTCTGTAGGCTTTTCAG 23
Db 18431 AAGACCTCTGTGTGGCTTTTCAG 18452

RESULT 34
US-09-949-016-12837
; Sequence 12837, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12837
; LENGTH: 148405
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12837

Query Match      68.8%; Score 17.2; DB 3; Length 148405;
Best Local Similarity 86.4%; Pred. No. 3.4e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAAGCTCTGTAGGCTTTTCAG 23
Db 18431 AAGACCTCTGTGTGGCTTTTCAG 18452

RESULT 35
US-09-676-610B-24/C
; Sequence 24, Application US/09676610B
; Patent No. 644465
; GENERAL INFORMATION:
; APPLICANT: C. Frank Bennett
; APPLICANT: Jacqueline Wyatt
; APPLICANT: Susan M. Freier
```

```
; TITLE OF INVENTION: OLIGONUCLEOTIDE INHIBITION OF HER-1 EXPRESSION
; FILE REFERENCE: RTS-0138
; CURRENT APPLICATION NUMBER: US/09/676,610B
; CURRENT FILING DATE: 2000-09-29
; NUMBER OF SEQ ID NOS: 182
; SEQ ID NO 24
; LENGTH: 169998
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: exon
; LOCATION: (1208)...(1472)
; NAME/KEY: intron
; LOCATION: (1473)...(124390)
; NAME/KEY: exon
; LOCATION: (124391)...(124544)
; NAME/KEY: intron
; LOCATION: (124545)...(125409)
; NAME/KEY: exon
; LOCATION: (125410)...(125595)
; NAME/KEY: intron
; LOCATION: (125596)...(128711)
; NAME/KEY: exon
; LOCATION: (128712)...(128848)
; NAME/KEY: intron
; LOCATION: (128849)...(133400)
; NAME/KEY: exon
; LOCATION: (133401)...(133469)
; NAME/KEY: intron
; LOCATION: (133470)...(134652)
; NAME/KEY: exon
; LOCATION: (134653)...(134773)
; NAME/KEY: intron
; LOCATION: (134774)...(136116)
; NAME/KEY: exon
; LOCATION: (136117)...(136261)
; NAME/KEY: intron
; LOCATION: (136262)...(137936)
; NAME/KEY: exon
; LOCATION: (137937)...(138053)
; NAME/KEY: intron
; LOCATION: (138054)...(138637)
; NAME/KEY: exon
; LOCATION: (138638)...(138766)
; NAME/KEY: intron
; LOCATION: (138767)...(138864)
; NAME/KEY: exon
; LOCATION: (138865)...(138940)
; NAME/KEY: intron
; LOCATION: (138941)...(139765)
; NAME/KEY: exon
; LOCATION: (139766)...(139860)
; NAME/KEY: intron
; LOCATION: (139861)...(142245)
; NAME/KEY: exon
; LOCATION: (142246)...(142445)
; NAME/KEY: intron
; LOCATION: (142446)...(143605)
; NAME/KEY: exon
; LOCATION: (143606)...(143738)
; NAME/KEY: intron
; LOCATION: (143739)...(145838)
; NAME/KEY: exon
; LOCATION: (145839)...(145931)
; NAME/KEY: intron
; LOCATION: (145932)...(147385)
; NAME/KEY: exon
; LOCATION: (147386)...(147544)
; NAME/KEY: intron
; LOCATION: (147545)...(153274)
; NAME/KEY: exon
; LOCATION: (153275)...(153321)
; NAME/KEY: intron
```

```
; LOCATION: (153322)...(155088)
; NAME/KEY: exon
; LOCATION: (155089)...(155231)
; NAME/KEY: intron
; LOCATION: (155232)...(156025)
; NAME/KEY: exon
; LOCATION: (156026)...(156151)
; NAME/KEY: intron
; LOCATION: (156152)...(156826)
; NAME/KEY: exon
; LOCATION: (156827)...(156928)
; NAME/KEY: intron
; LOCATION: (156929)...(163399)
; NAME/KEY: exon
; LOCATION: (163400)...(163586)
; US-09-676-610B-24
Query Match 68.8%; Score 17.2; DB 3; Length 169998;
Best Local Similarity 86.4%; Pred. No. 3.5e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 2 AAAAACTCTGTAGGCTTTTCAG 23
|||||
Db 10676 AAAATCTCTGAAAGCTTTCAG 10655
|||||
RESULT 36
US-09-949-016-15841
; Sequence 15841, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15841
; LENGTH: 177251
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-15841
Query Match 68.8%; Score 17.2; DB 3; Length 177251;
Best Local Similarity 86.4%; Pred. No. 3.5e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 AAAAACTCTGTAGGCTTTTC 22
|||||
Db 79470 AAAAACTCTGTAAACATTC 79491
|||||
RESULT 37
US-09-877-177A-10/c
; Sequence 10, Application US/09877177A
; Patent No. 6582919
; GENERAL INFORMATION:
; APPLICANT: K. Danenberg
; TITLE OF INVENTION: Method of determining Epidermal Growth
; Factor Receptor and HER2-Neu Gene Expression
; TITLE OF INVENTION: and Correlation of Levels Thereof With Survival
; FILE REFERENCE: 11220/120
; CURRENT APPLICATION NUMBER: US/09/877,177A
; CURRENT FILING DATE: 2001-06-11
; NUMBER OF SEQ ID NOS: 11
```

```
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 197496
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-877-177A-10

Query Match      68.8%; Score 17.2; DB 3; Length 197496;
Best Local Similarity 86.4%; Pred. No. 3.6e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAACTCCTGTAGGCTTTTCAG 23
    ||||| ||||| ||||| ||||| |||||
Db 18676 AAAATCTCTGAAGCTTTTCAG 18655

RESULT 38
US-09-949-016-133546/c
; Sequence 133546, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 133546
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-133546

Query Match      68.0%; Score 17; DB 3; Length 601;
Best Local Similarity 80.0%; Pred. No. 1e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAACTCCTGTAGGCTTTTCAGTG 25
    ||||| ||||| ||||| ||||| |||||
Db 428 AAAAAACAATTGTAGCTTTTCATG 404

RESULT 39
US-09-949-016-133547/c
; Sequence 133547, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 133547
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
```

```
US-09-949-016-133547

Query Match      68.0%; Score 17; DB 3; Length 601;
Best Local Similarity 80.0%; Pred. No. 1e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAACTCCTGTAGGCTTTTCAGTG 25
    ||||| ||||| ||||| ||||| |||||
Db 406 AAAAAACAATTGTAGCTTTTCATG 382

RESULT 40
US-09-270-767-982
; Sequence 982, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 982
; LENGTH: 703
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-982

Query Match      68.0%; Score 17; DB 3; Length 703;
Best Local Similarity 80.0%; Pred. No. 1.1e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAACTCCTGTAGGCTTTTCAGTG 25
    ||||| ||||| ||||| ||||| |||||
Db 560 AAAAAAATGCTTTAAGCTTTTCATTG 584

RESULT 41
US-09-270-767-16264
; Sequence 16264, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16264
; LENGTH: 703
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-16264

Query Match      68.0%; Score 17; DB 3; Length 703;
Best Local Similarity 80.0%; Pred. No. 1.1e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAACTCCTGTAGGCTTTTCAGTG 25
    ||||| ||||| ||||| ||||| |||||
Db 560 AAAAAAATGCTTTAAGCTTTTCATTG 584

RESULT 42
US-09-134-000C-625/c
; Sequence 625, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
```

```
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 625
; LENGTH: 1095
; TYPE: DNA
; ORGANISM: Enterococcus faecalis
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1051)..(1051)
; OTHER INFORMATION: Nucleotide 1051 is "n" wherein "n" = any nucleotide.
US-09-134-000C-625

Query Match      68.0%; Score 17; DB 3; Length 1095;
Best Local Similarity 80.0%; Pred. No. 1.2e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAACTCTGTAGGCTTTCAGTG 25
   ||||| ||||| ||||| ||||| ||
Db 990 AAAAACTCTGTAGGCTTTCATG 966

RESULT 43
US-09-799-451-566/c
; Sequence 566, Application US/09799451
; Patent No. 6783969
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Ryle
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Zhang, Jie
; APPLICANT: Xue, Aidong J.
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Ma, Yunqing
; APPLICANT: Yamazaki, Victoria
; APPLICANT: Chen, Rui-hong
; APPLICANT: Wang, Zhiwei
; APPLICANT: Wang, Dunrui
; APPLICANT: Yang, Yonghong
; APPLICANT: Wehrman, Tom
; APPLICANT: Ghosh, Reena
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6783969el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 803
; CURRENT APPLICATION NUMBER: US/09/799,451
; CURRENT FILING DATE: 2001-03-05
; NUMBER OF SEQ ID NOS: 948
; SOFTWARE: Pt_Fl_genes Version 2.0
; SEQ ID NO 566
; LENGTH: 1939
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (145)..(678)
US-09-799-451-566

Query Match      68.0%; Score 17; DB 3; Length 1939;
Best Local Similarity 80.0%; Pred. No. 1.4e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAACTCTGTAGGCTTTCAGTG 25
   ||||| ||||| ||||| ||||| ||
Db 1185 AAAAACTCTATAGGCTTAAAGTG 1161

; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 625
; LENGTH: 1095
; TYPE: DNA
; ORGANISM: Enterococcus faecalis
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1051)..(1051)
; OTHER INFORMATION: Nucleotide 1051 is "n" wherein "n" = any nucleotide.
US-09-134-000C-625

Query Match      68.0%; Score 17; DB 3; Length 1095;
Best Local Similarity 80.0%; Pred. No. 1.2e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAACTCTGTAGGCTTTCAGTG 25
   ||||| ||||| ||||| ||||| ||
Db 990 AAAAACTCTGTAGGCTTTCATG 966

RESULT 43
US-09-799-451-566/c
; Sequence 566, Application US/09799451
; Patent No. 6783969
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Ryle
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Zhang, Jie
; APPLICANT: Xue, Aidong J.
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Ma, Yunqing
; APPLICANT: Yamazaki, Victoria
; APPLICANT: Chen, Rui-hong
; APPLICANT: Wang, Zhiwei
; APPLICANT: Wang, Dunrui
; APPLICANT: Yang, Yonghong
; APPLICANT: Wehrman, Tom
; APPLICANT: Ghosh, Reena
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6783969el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 803
; CURRENT APPLICATION NUMBER: US/09/799,451
; CURRENT FILING DATE: 2001-03-05
; NUMBER OF SEQ ID NOS: 948
; SOFTWARE: Pt_Fl_genes Version 2.0
; SEQ ID NO 566
; LENGTH: 1939
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (145)..(678)
US-09-799-451-566

Query Match      68.0%; Score 17; DB 3; Length 1939;
Best Local Similarity 80.0%; Pred. No. 1.4e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAACTCTGTAGGCTTTCAGTG 25
   ||||| ||||| ||||| ||||| ||
Db 1185 AAAAACTCTATAGGCTTAAAGTG 1161

RESULT 44
US-09-949-016-1046/c
; Sequence 1046, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1046
; LENGTH: 2905
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-1046

Query Match      68.0%; Score 17; DB 3; Length 2905;
Best Local Similarity 80.0%; Pred. No. 1.5e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAACTCTGTAGGCTTTCAGTG 25
   ||||| ||||| ||||| ||||| ||
Db 586 AAAAACTCTGTGGAGTTCAGTG 562

RESULT 45
US-09-949-016-4827/c
; Sequence 4827, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4827
; LENGTH: 3416
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-4827

Query Match      68.0%; Score 17; DB 3; Length 3416;
Best Local Similarity 80.0%; Pred. No. 1.6e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAACTCTGTAGGCTTTCAGTG 25
   ||||| ||||| ||||| ||||| ||
Db 586 AAAAACTCTGTGGAGTTCAGTG 562

RESULT 46
US-09-949-016-4828/c
; Sequence 4828, Application US/09949016
; Patent No. 6812339
```

```
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4828
; LENGTH: 3416
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-4828

Query Match          68.0%; Score 17; DB 3; Length 3416;
Best Local Similarity 80.0%; Pred. No. 1.6e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAACTCCTGTAGGCTTTCAGTG 25
   ||||| ||| || |||||
Db 586 AAAAACTCTGCTGGAGTTCAGTG 562

RESULT 47
US-09-949-016-4829/c
; Sequence 4829, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4829
; LENGTH: 3490
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-4829

Query Match          68.0%; Score 17; DB 3; Length 3490;
Best Local Similarity 80.0%; Pred. No. 1.6e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAACTCCTGTAGGCTTTCAGTG 25
   ||||| ||| || |||||
Db 586 AAAAACTCTGCTGGAGTTCAGTG 562

RESULT 48
US-09-949-016-4830/c
; Sequence 4830, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
```

```
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4830
; LENGTH: 3490
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-4830

Query Match          68.0%; Score 17; DB 3; Length 3490;
Best Local Similarity 80.0%; Pred. No. 1.6e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAACTCCTGTAGGCTTTCAGTG 25
   ||||| ||| || |||||
Db 586 AAAAACTCTGCTGGAGTTCAGTG 562

RESULT 49
US-09-949-016-15361
; Sequence 15361, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15361
; LENGTH: 17132
; TYPE: DNA
; ORGANISM: Human
; NAME/KEY: misc feature
; LOCATION: (1)...(17132)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15361

Query Match          68.0%; Score 17; DB 3; Length 17132;
Best Local Similarity 80.0%; Pred. No. 2.5e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAACTCCTGTAGGCTTTCAGTG 25
   ||||| ||| || |||||
Db 6872 AAAAAATGCTGTAGTGGCTTTCAGAG 6896

RESULT 50
US-09-949-016-12438
; Sequence 12438, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
```

Mon Feb 6 12:23:14 2006

;
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: PastSeq for Windows Version 4.0
; SEQ ID NO 12438
; LENGTH: 24150
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(24150)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12438

Query Match 68.0%; Score 17; DB 3; Length 24150;
Best Local Similarity 80.0%; Pred. No. 2.7e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAACTCCTGTAGGCTTTCAGTG 25
||||| ||||| ||||| ||||| |||||
DB 6872 AAAAAATGCTGATGGCTTTCAGAG 6896

Search completed: February 3, 2006, 16:32:30
Job time : 92.6667 secs

HIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.
OM nucleic - nucleic search, using sw model
Run on: February 3, 2006, 14:37:28 ; Search time 365.556 Seconds
(without alignments)
565.535 Million cell updates/sec

Title: US-10-719-900-20
Perfect score: 25
Sequence: 1 aaaaactctgtagcttcagtg 25
Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0
Searched: 9793542 seqs, 4134689005 residues
Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 150 summaries

Database : ; Published Applications NA Main:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:
3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:
4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:
6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:
7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:
8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:
9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:
Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES				Description	
Result No.	Score	Query Match	Length DB ID		
1	25	100.0	25 8	US-10-719-900-20	Sequence 20, Appl
C 2	25	100.0	74868 5	US-10-175-523-67	Sequence 67, Appl
C 3	25	100.0	74868 10	US-11-099-266-67	Sequence 67, Appl
4	23.4	93.6	25 8	US-10-719-900-19	Sequence 19, Appl
C 5	19.8	79.2	334 8	US-10-425-115-33079	Sequence 33079, A
C 6	19.8	79.2	2243 8	US-10-788-792-104	Sequence 104, App
C 7	19.2	76.8	258 7	US-10-242-535A-7089	Sequence 7089, App
C 8	19.2	76.8	258 7	US-10-085-783A-7089	Sequence 7089, App
C 9	19.2	76.8	289 7	US-10-242-535A-38667	Sequence 38667, A
C 10	19.2	76.8	289 7	US-10-085-783A-38667	Sequence 38667, A
C 11	19.2	76.8	439 6	US-10-125-968-712	Sequence 712, App
C 12	19.2	76.8	546 5	US-10-027-632-199786	Sequence 199786, A
C 13	19.2	76.8	546 6	US-10-027-632-199786	Sequence 199786, A
C 14	19.2	76.8	562 3	US-09-864-761-13669	Sequence 13669, A
C 15	19.2	76.8	693 4	US-09-925-065A-81129	Sequence 81129, A
C 16	19.2	76.8	1400 9	US-10-956-157-6011	Sequence 6011, App
C 17	19.2	76.8	2527 6	US-10-108-260A-731	Sequence 731, App
C 18	19.2	76.8	2587 5	US-10-146-473-6	Sequence 6, Appli
C 19	19.2	76.8	2587 5	US-10-717-597-125	Sequence 125, App
C 20	19.2	76.8	2587 9	US-10-956-157-776	Sequence 776, App
C 21	19.2	76.8	2587 9	US-10-981-277-8	Sequence 8, Appli
C 22	19.2	76.8	3886 6	US-10-252-157-57	Sequence 57, Appl
C 23	19.2	76.8	5569 3	US-09-814-353-20129	Sequence 20129, A

C	24	19.2	76.8	5570	8	US-10-357-930-24829	Sequence 24829, A
	25	19.2	76.8	150525	9	US-10-981-277-36	Sequence 36, Appl
	26	19.2	76.8	166536	9	US-10-981-277-35	Sequence 35, Appl
	27	18.6	74.4	594	5	US-10-027-632-129877	Sequence 129877, A
	28	18.6	74.4	594	5	US-10-027-632-129878	Sequence 129878, A
	29	18.6	74.4	594	6	US-10-027-632-129877	Sequence 129877, A
	30	18.6	74.4	594	6	US-10-027-632-129878	Sequence 129878, A
	31	18.6	74.4	823	5	US-10-027-632-133274	Sequence 133274, A
	32	18.6	74.4	823	5	US-10-027-632-133275	Sequence 133275, A
	33	18.6	74.4	823	6	US-10-027-632-133274	Sequence 133274, A
	34	18.6	74.4	823	6	US-10-027-632-133275	Sequence 133275, A
	35	18.6	74.4	39982	6	US-10-369-626-3	Sequence 3, Appli
	36	18.6	74.4	39982	8	US-10-900-420-3	Sequence 3, Appli
C	37	18.2	72.8	317	8	US-10-425-115-172275	Sequence 172275, A
C	38	18.2	72.8	351	8	US-10-425-115-33082	Sequence 33082, A
	39	18.2	72.8	429	4	US-09-925-065A-814695	Sequence 814695, A
	40	18.2	72.8	600	9	US-10-972-079-59759	Sequence 59759, A
	41	18.2	72.8	638	4	US-09-925-065A-450957	Sequence 450957, A
	42	18.2	72.8	660	4	US-09-925-065A-554692	Sequence 554692, A
	43	18.2	72.8	834	4	US-09-925-065A-681797	Sequence 681797, A
	44	18.2	72.8	919	5	US-10-027-632-121229	Sequence 121229, A
	45	18.2	72.8	919	5	US-10-027-632-121230	Sequence 121230, A
	46	18.2	72.8	919	6	US-10-027-632-121229	Sequence 121229, A
	47	18.2	72.8	919	6	US-10-027-632-121230	Sequence 121230, A
C	48	18.2	72.8	1136	5	US-10-027-632-117480	Sequence 117480, A
C	49	18.2	72.8	1136	6	US-10-027-632-117480	Sequence 117480, A
C	50	18.2	72.8	1175	7	US-10-424-599-42397	Sequence 42397, A
C	51	18.2	72.8	1985	5	US-10-027-632-97548	Sequence 97548, A
C	52	18.2	72.8	1985	5	US-10-027-632-99669	Sequence 99669, A
C	53	18.2	72.8	1985	6	US-10-027-632-97548	Sequence 97548, A
C	54	18.2	72.8	1985	6	US-10-027-632-99669	Sequence 99669, A
C	55	18.2	72.8	2906	5	US-10-027-632-249510	Sequence 249510, A
C	56	18.2	72.8	2906	6	US-10-027-632-249510	Sequence 249510, A
C	57	18.2	72.8	112341	7	US-10-322-281-656	Sequence 656, App
C	58	17.8	71.2	429	4	US-09-925-065A-277450	Sequence 277450, A
C	59	17.8	71.2	564	4	US-09-925-065A-781841	Sequence 781841, A
C	60	17.8	71.2	760	9	US-10-756-149-3366	Sequence 3366, Ap
C	61	17.8	71.2	2281	4	US-09-925-065A-722928	Sequence 722928, A
C	62	17.8	71.2	2281	4	US-09-925-065A-722929	Sequence 722929, A
C	63	17.8	71.2	1163020	7	US-10-398-221-10	Sequence 10, Appl
C	64	17.8	71.2	301208	7	US-10-398-221-2058	Sequence 2058, Ap
	65	17.6	70.4	436	4	US-09-925-065A-335284	Sequence 335284, A
	66	17.6	70.4	465	4	US-09-925-065A-478124	Sequence 478124, A
	67	17.6	70.4	465	4	US-09-925-065A-478125	Sequence 478125, A
C	68	17.6	70.4	490	4	US-09-925-065A-601469	Sequence 601469, A
C	69	17.6	70.4	533	7	US-10-437-963-99308	Sequence 99308, A
C	70	17.6	70.4	600	9	US-10-972-079-34352	Sequence 34352, A
	71	17.6	70.4	647	5	US-10-027-632-230481	Sequence 230481, A
	72	17.6	70.4	647	6	US-10-027-632-230481	Sequence 230481, A
C	73	17.6	70.4	671	4	US-09-925-065A-770924	Sequence 770924, A
	74	17.6	70.4	851	7	US-10-437-963-14931	Sequence 14931, A
C	75	17.6	70.4	925	6	US-10-369-493-32578	Sequence 32578, A
C	76	17.6	70.4	1303	7	US-10-425-114-4266	Sequence 4266, Ap
	77	17.6	70.4	1368	10	US-11-097-143-14957	Sequence 14957, A
	78	17.6	70.4	1563	3	US-09-160-116-1	Sequence 1, Appli
	79	17.6	70.4	1590	4	US-09-925-065A-53317	Sequence 53317, A
	80	17.6	70.4	1590	4	US-09-925-065A-53318	Sequence 53318, A
	81	17.6	70.4	1660	7	US-10-437-963-14930	Sequence 14930, A
	82	17.6	70.4	1867	3	US-09-879-957-13	Sequence 13, Appl
	83	17.6	70.4	1867	7	US-10-807-856-13	Sequence 13, Appl
	84	17.6	70.4	1972	7	US-10-437-963-14927	Sequence 14927, A
C	85	17.6	70.4	5010	10	US-11-097-143-14956	Sequence 14956, A
C	86	17.6	70.4	5904	10	US-11-097-143-25043	Sequence 25043, A
C	87	17.6	70.4	8685	10	US-11-097-143-25042	Sequence 25042, A
C	88	17.6	70.4	90468	6	US-10-085-117-160	Sequence 160, App
	89	17.6	70.4	99232	8	US-10-331-053-16	Sequence 16, Appl
	90	17.6	70.4	99588	5	US-10-087-192-1780	Sequence 1780, Ap
	91	17.6	70.4	227968	8	US-10-723-860-1357	Sequence 1357, Ap
C	92	17.4	69.6	429	4	US-09-925-065A-277451	Sequence 277451, A
C	93	17.4	69.6	591	4	US-09-925-065A-659385	Sequence 659385, A
C	94	17.4	69.6	591	4	US-09-925-065A-659386	Sequence 659386, A
C	95	17.4	69.6	609	4	US-09-925-065A-355809	Sequence 355809, A
C	96	17.4	69.6	616	5	US-10-027-632-271180	Sequence 271180, A

c 97	17.4	69.6	616	5	US-10-027-632-271181	Sequence 271181,
c 98	17.4	69.6	616	6	US-10-027-632-271180	Sequence 271180,
c 99	17.4	69.6	616	6	US-10-027-632-271181	Sequence 271181,
c 100	17.4	69.6	635	4	US-09-925-065A-522160	Sequence 522160,
c 101	17.2	68.8	25	7	US-10-719-900-20	Sequence 536957,
c 102	17.2	68.8	356	7	US-10-469-285-595	Sequence 595,
c 103	17.2	68.8	426	3	US-09-918-995-3989	Sequence 3989, Ap
c 104	17.2	68.8	524	8	US-10-425-115-72882	Sequence 72882, A
c 105	17.2	68.8	576	9	US-10-779-543-16050	Sequence 16050, A
c 106	17.2	68.8	584	4	US-09-925-065A-516746	Sequence 516746,
c 107	17.2	68.8	599	9	US-10-972-079-50972	Sequence 50972, A
c 108	17.2	68.8	600	9	US-10-972-079-50973	Sequence 50973, A
c 109	17.2	68.8	806	4	US-09-925-065A-7814	Sequence 7814, Ap
c 110	17.2	68.8	883	5	US-10-027-632-28833	Sequence 28833, A
c 111	17.2	68.8	883	6	US-10-027-632-28833	Sequence 28833, A
c 112	17.2	68.8	1093	4	US-09-925-065A-287411	Sequence 287411,
c 113	17.2	68.8	1315	5	US-10-106-698-891	Sequence 891, App
c 114	17.2	68.8	1348	8	US-10-723-860-3086	Sequence 3086, Ap
c 115	17.2	68.8	1579	5	US-10-027-632-264994	Sequence 264994,
c 116	17.2	68.8	1579	6	US-10-027-632-264994	Sequence 264994,
c 117	17.2	68.8	1857	7	US-10-282-122A-17442	Sequence 17442, A
c 118	17.2	68.8	1860	8	US-10-370-715B-349	Sequence 349, App
c 119	17.2	68.8	1956	8	US-10-723-860-7107	Sequence 7107, Ap
c 120	17.2	68.8	1971	5	US-10-037-270-960	Sequence 960, App
c 121	17.2	68.8	1971	6	US-10-117-722-960	Sequence 960, App
c 122	17.2	68.8	2061	9	US-10-122-851-960	Sequence 960, App
c 123	17.2	68.8	2061	9	US-10-450-763-15246	Sequence 15246, A
c 124	17.2	68.8	2022	3	US-09-764-881-84	Sequence 84, Appl
c 125	17.2	68.8	2022	3	US-09-764-853-428	Sequence 428, App
c 126	17.2	68.8	2022	4	US-09-764-881-84	Sequence 84, Appl
c 127	17.2	68.8	2022	6	US-10-242-747-84	Sequence 84, Appl
c 128	17.2	68.8	2251	4	US-09-925-065A-84940	Sequence 84940, A
c 129	17.2	68.8	3376	3	US-09-834-975-775	Sequence 775, App
c 130	17.2	68.8	3914	6	US-10-104-047-850	Sequence 850, App
c 131	17.2	68.8	4263	6	US-10-406-209-7	Sequence 7, Appli
c 132	17.2	68.8	6138	7	US-10-275-287-4	Sequence 4, Appli
c 133	17.2	68.8	6141	7	US-10-275-287-3	Sequence 3, Appli
c 134	17.2	68.8	6586	5	US-10-119-926-41	Sequence 41, Appl
c 135	17.2	68.8	6586	6	US-10-291-172-86	Sequence 86, Appl
c 136	17.2	68.8	6586	7	US-10-221-278-86	Sequence 86, Appl
c 137	17.2	68.8	7646	7	US-10-275-287-1	Sequence 1, Appli
c 138	17.2	68.8	13927	8	US-10-719-993-7032	Sequence 7032, Ap
c 139	17.2	68.8	16570	7	US-10-275-287-2	Sequence 2, Appli
c 140	17.2	68.8	24069	8	US-10-719-993-6828	Sequence 6828, Ap
c 141	17.2	68.8	25473	8	US-10-719-993-7034	Sequence 7034, Ap
c 142	17.2	68.8	68123	7	US-10-741-601-5697	Sequence 5697, Ap
c 143	17.2	68.8	68123	8	US-10-741-601-5697	Sequence 17774, A
c 144	17.2	68.8	108359	5	US-10-191-807-3	Sequence 3, Appli
c 145	17.2	68.8	108359	9	US-10-959-243-3	Sequence 3, Appli
c 146	17.2	68.8	169998	6	US-10-380-931-24	Sequence 24, Appl
c 147	17.2	68.8	197496	3	US-09-877-177-10	Sequence 10, Appl
c 148	17.2	68.8	197496	6	US-10-426-836-10	Sequence 10, Appl
c 149	17.2	68.8	200620	9	US-10-704-513-4	Sequence 4, Appli
c 150	17.2	68.8	209083	9	US-10-461-862-74	Sequence 74, Appl

SEQ ID NO 20

LENGTH: 25

TYPE: DNA

ORGANISM: Mus musculus

US-10-719-900-20

Query Match 100.0%; Score 25; DB 8; Length 25;

Best Local Similarity 100.0%; Pred. No. 0.039;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAACTCCTGCTAGGCTTTTCAGTG 25

|||||

DB 1 AAAAACTCCTGCTAGGCTTTTCAGTG 25

|||||

RESULT 2

US-10-175-523-67/c

; Sequence 67, Application US/10175523

; Publication No. US20030096264A1

; GENERAL INFORMATION:

; APPLICANT: Brockman, Jeffrey

; APPLICANT: Evans, David

; APPLICANT: Hook, Derek

; APPLICANT: Klimczak, Leszek

; APPLICANT: Laeng, Pascal

; APPLICANT: Palfreyman, Michael

; APPLICANT: Rajan, Prithi

; TITLE OF INVENTION: MULTI-PARAMETER HIGH THROUGHPUT SCREENING ASSAYS (MPHTS)

; FILE REFERENCE: 3235/1J795-US3

; CURRENT APPLICATION NUMBER: US/10/175,523

; CURRENT FILING DATE: 2002-06-18

; PRIOR APPLICATION NUMBER: US 60/299,151

; PRIOR FILING DATE: 2001-06-18

; PRIOR APPLICATION NUMBER: US 60/317,828

; PRIOR FILING DATE: 2001-09-07

; PRIOR APPLICATION NUMBER: US 60/325,150

; PRIOR FILING DATE: 2001-09-25

; PRIOR APPLICATION NUMBER: US 60/333,047

; PRIOR FILING DATE: 2001-11-14

; PRIOR APPLICATION NUMBER: US 60/349,936

; PRIOR FILING DATE: 2002-01-18

; PRIOR APPLICATION NUMBER: US 60/361,834

; PRIOR FILING DATE: 2002-03-04

; NUMBER OF SEQ ID NOS: 197

; SOFTWARE: PatentIn version 3.1

SEQ ID NO 67

LENGTH: 74868

TYPE: DNA

ORGANISM: Mus musculus

US-10-175-523-67

Query Match 100.0%; Score 25; DB 5; Length 74868;

Best Local Similarity 100.0%; Pred. No. 0.12;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAACTCCTGCTAGGCTTTTCAGTG 25

|||||

DB 33482 AAAAACTCCTGCTAGGCTTTTCAGTG 33458

|||||

RESULT 3

US-11-099-266-67/c

; Sequence 67, Application US/11099266

; Publication No. US20050181433A1

; GENERAL INFORMATION:

; APPLICANT: Brockman, Jeffrey

; APPLICANT: Evans, David

; APPLICANT: Hook, Derek

; APPLICANT: Klimczak, Leszek

; APPLICANT: Laeng, Pascal

; APPLICANT: Palfreyman, Michael

; APPLICANT: Rajan, Prithi

; TITLE OF INVENTION: MULTI-PARAMETER HIGH THROUGHPUT SCREENING ASSAYS (MPHTS)

; FILE REFERENCE: 3528.1

; CURRENT APPLICATION NUMBER: US/10/719,900

; PRIOR FILING DATE: 2003-11-20

; PRIOR APPLICATION NUMBER: 60/427,808

; PRIOR FILING DATE: 2002 11 20

; NUMBER OF SEQ ID NOS: 982914

; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1

```
; FILE REFERENCE: 03235/100J795-US4
; CURRENT APPLICATION NUMBER: US/11/099,266
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: US 10/175,523
; PRIOR FILING DATE: 2002-06-18
; PRIOR APPLICATION NUMBER: US 60/299,151
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: US 60/317,828
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: US 60/325,150
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/333,047
; PRIOR FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: US 60/349,936
; PRIOR FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: US 60/361,834
; PRIOR FILING DATE: 2002-03-04
; NUMBER OF SEQ ID NOS: 197
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 67
; LENGTH: 74868
; TYPE: DNA
; ORGANISM: Mus musculus
US-11-099-266-67
```

```
Query Match 100.0%; Score 25; DB 10; Length 74868;
Best Local Similarity 100.0%; Pred. No. 0.12; Indels 0; Gaps 0;
Matches 25; Conservative 0; Mismatches 0;
```

```
QY 1 AAAAACTCCTGTAGGCTTTTCAGTG 25
|||||
DB 33482 AAAAACTCCTGTAGGCTTTTCAGTG 33458
```

RESULT 4

```
US-10-719-900-19
; Sequence 19, Application US/10719900
; Publication No. US20050026164A1
; GENERAL INFORMATION:
; APPLICANT: Xue Mei Zhou
; TITLE OF INVENTION: Methods of Genetic Analysis of Mouse
; FILE REFERENCE: 3528.1
; CURRENT APPLICATION NUMBER: US/10/719,900
; CURRENT FILING DATE: 2003-11-20
; PRIOR APPLICATION NUMBER: 60/427,808
; PRIOR FILING DATE: 2002 11 20
; NUMBER OF SEQ ID NOS: 982914
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 19
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-719-900-19
```

```
Query Match 93.6%; Score 23.4; DB 8; Length 25;
Best Local Similarity 96.0%; Pred. No. 0.24; Indels 1; Gaps 0;
Matches 24; Conservative 0; Mismatches 0;
```

```
QY 1 AAAAACTCCTGTAGGCTTTTCAGTG 25
|||||
DB 1 AAAAACTCCTGTAGGCTTTTCAGTG 25
```

RESULT 5

```
US-10-425-115-33079/c
; Sequence 33079, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
```

```
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 33079
; LENGTH: 334
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_130178C.1
US-10-425-115-33079
```

```
Query Match 79.2%; Score 19.8; DB 8; Length 334;
Best Local Similarity 91.3%; Pred. No. 21;
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 1 AAAAACTCCTGTAGGCTTTTCAG 23
|||||
DB 101 AAAAACTCCTGTAGGCTGTGAG 79
```

RESULT 6

```
US-10-788-792-104/c
; Sequence 104, Application US/10788792
; Publication No. US20040191819A1
; GENERAL INFORMATION:
; APPLICANT: Bayer Pharmaceuticals Corporation
; APPLICANT: Eyeleight, Deepa
; APPLICANT: Bigwood, Douglas
; TITLE OF INVENTION: EXPRESSION PROFILES FOR BREAST CANCER AND METHODS OF USE
; FILE REFERENCE: 5152
; CURRENT APPLICATION NUMBER: US/10/788,792
; CURRENT FILING DATE: 2004-02-27
; PRIOR APPLICATION NUMBER: US 60/450,655
; PRIOR FILING DATE: 2003-02-28
; NUMBER OF SEQ ID NOS: 254
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 104
; LENGTH: 2243
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (102)..(102)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (156)..(156)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (493)..(493)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (498)..(498)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (506)..(506)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (535)..(535)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (539)..(539)
```



```
; NAME/KEY: misc_feature
; LOCATION: (1890)..(1890)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature

Query Match          79.2%; Score 19.8; DB 8; Length 2243;
Best Local Similarity 87.5%; Pred. No. 28;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAACTCCTGTAGGCTTTTCAGT 24
   ||||| ||||| ||||| ||||| |||||
Db 2095 AAAAAATTCCTGTAGGTTTTCANT 2072

RESULT 7
US-10-242-535A-7089/c
; Sequence 7089, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liaw, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2005
; CURRENT APPLICATION NUMBER: US/10/242,535A
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 10/085,783
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 7089
; LENGTH: 258
; TYPE: DNA
; ORGANISM: Human
US-10-242-535A-7089

Query Match          76.8%; Score 19.2; DB 7; Length 258;
Best Local Similarity 87.5%; Pred. No. 40;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAACTCCTGTAGGCTTTTCAGT 24
   ||||| ||||| ||||| ||||| |||||
Db 136 AAAAAATTCCTGTAGGTTTTCATT 113

RESULT 8
US-10-085-783A-7089/c
; Sequence 7089, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liaw, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2002
; CURRENT APPLICATION NUMBER: US/10/085,783A
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 7089
; LENGTH: 258
; TYPE: DNA
; ORGANISM: Human
US-10-085-783A-7089

; ORGANISM: Human
US-10-085-783A-7089

Query Match          76.8%; Score 19.2; DB 7; Length 258;
Best Local Similarity 87.5%; Pred. No. 40;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAACTCCTGTAGGCTTTTCAGT 24
   ||||| ||||| ||||| ||||| |||||
Db 166 AAAAAATTCCTGTAGGTTTTCATT 143

RESULT 10
US-10-085-783A-38667/c
; Sequence 38667, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liaw, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2002
; CURRENT APPLICATION NUMBER: US/10/085,783A
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 38667
; LENGTH: 289
; TYPE: DNA
; ORGANISM: Human
US-10-242-535A-38667
```

Query Match 76.8%; Score 19.2; DB 7; Length 289;
Best Local Similarity 87.5%; Pred. No. 41;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAACTCTGTAGGTTTCAGT 24
||||| ||||| ||||| ||||| |||||
Db 166 AAAAAATTCCTGTAGGTTTCATT 143

RESULT 11
US-10-125-968-712/c
; Sequence 712, Application US/10125968
; Publication No. US20030215805A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Palermo, Adam
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; APPLICANT: Elias, Josh
; APPLICANT: Mertens, Maureen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-032
; CURRENT APPLICATION NUMBER: US/10/125,968
; CURRENT FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: US 60/285,163
; PRIOR FILING DATE: 2001-04-20
; NUMBER OF SEQ ID NOS: 1417
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 712
; LENGTH: 439
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 290..291, 295, 297, 302, 308, 319, 340, 395, 408, 416, 423,
; LOCATION: 424
; OTHER INFORMATION: n = A,T,C or G
US-10-125-968-712

Query Match 76.8%; Score 19.2; DB 6; Length 439;
Best Local Similarity 87.5%; Pred. No. 43;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAACTCTGTAGGTTTCAGT 24
||||| ||||| ||||| ||||| |||||
Db 166 AAAAAATTCCTGTAGGTTTCATT 143

RESULT 12
US-10-027-632-199786/c
; Sequence 199786, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28

; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 199786
; LENGTH: 546
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-199786

Query Match 76.8%; Score 19.2; DB 5; Length 546;
Best Local Similarity 87.5%; Pred. No. 45;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAACTCTGTAGGTTTCAGT 24
||||| ||||| ||||| ||||| |||||
Db 339 AAAAAATTCCTGTAGGTTTCATT 316

RESULT 13
US-10-027-632-199786/c
; Sequence 199786, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 199786
; LENGTH: 546
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-199786

Query Match 76.8%; Score 19.2; DB 6; Length 546;
Best Local Similarity 87.5%; Pred. No. 45;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAACTCTGTAGGTTTCAGT 24
||||| ||||| ||||| ||||| |||||
Db 339 AAAAAATTCCTGTAGGTTTCATT 316

RESULT 14
US-09-864-761-13669/c
; Sequence 13669, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1

; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
; SEQ ID NO 13669
; LENGTH: 562
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AP000356.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 5.3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.7
US-09-864-761-13669

Query Match 76.8%; Score 19.2; DB 3; Length 562;
Best Local Similarity 87.5%; Pred. No. 45;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 AAAAATCCTGTAGGCTTTTCAGTG 25
Db 29 AAAAATCCTGTAGGCTTTTCATTG 6

RESULT 15
US-09-925-065A-81129/c
; Sequence 81129, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925.065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24

; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 81129
; LENGTH: 693
; TYPE: DNA
; ORGANISM: Homo sapiens
; ORGANISM: Homo sapiens
US-09-925-065A-81129

Query Match 76.8%; Score 19.2; DB 4; Length 693;
Best Local Similarity 87.5%; Pred. No. 46;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 AAAAATCCTGTAGGCTTTTCAGTG 25
Db 199 AAAAATCCTGTAGGCTTTTCAGTG 176

RESULT 16
US-10-956-157-6011/c
; Sequence 6011, Application US/10956157
; Publication No. US20050118625A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William
; TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
; FILE REFERENCE: 031896-043000 (AM 101081)
; CURRENT APPLICATION NUMBER: US/10/956,157
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 319805
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6011
; LENGTH: 1400
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-956-157-6011

Query Match 76.8%; Score 19.2; DB 9; Length 1400;
Best Local Similarity 87.5%; Pred. No. 51;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AAAAATCCTGTAGGCTTTTCAGT 24
Db 1276 AAAAATCCTGTAGGCTTTTCATT 1253

RESULT 17
US-10-108-260A-731/c
; Sequence 731, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: NO. US20040005560A1el full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 731
; LENGTH: 2527
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-108-260A-731

Query Match 76.8%; Score 19.2; DB 6; Length 2527;

Best Local Similarity 87.5%; Pred. No. 55;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAAATCCTGTAGGTTTCAGT 24
|||||
DB 2405 AAAAAAATCCTGTAGGTTTCATT 2382

RESULT 18
US-10-146-473-6/c
; Sequence 6, Application US/10146473
; Publication No. US20030108988A1
; GENERAL INFORMATION:
; APPLICANT: Scanlan, Matthew
; APPLICANT: Gout, Ivan
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Gure, Ali
; APPLICANT: Chen, Yao-Tseng
; APPLICANT: Old, Lloyd
; TITLE OF INVENTION: Breast Cancer Antigens
; FILE REFERENCE: L00461/70130(JRV)
; CURRENT APPLICATION NUMBER: US/10/146,473
; CURRENT FILING DATE: 2002-05-15
; PRIOR APPLICATION NUMBER: US 60/291,150
; PRIOR FILING DATE: 2001-05-15
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 2587
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-146-473-6

Query Match 76.8%; Score 19.2; DB 5; Length 2587;
Best Local Similarity 87.5%; Pred. No. 56;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAAATCCTGTAGGTTTCAGT 24
|||||
DB 2463 AAAAAAATCCTGTAGGTTTCATT 2440

RESULT 19
US-10-717-597-125/c
; Sequence 125, Application US/10717597
; Publication No. US20040110221A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Burczynski, Michael E.
; APPLICANT: Twine, Natalie C.
; APPLICANT: Dornier, Andrew J.
; APPLICANT: Trepicchio, William L.
; APPLICANT: Slonim, Donna K.
; APPLICANT: Stover, Jennifer A.
; TITLE OF INVENTION: METHODS FOR DIAGNOSING RCC AND OTHER SOLID TUMORS
; FILE REFERENCE: AM01080L
; CURRENT APPLICATION NUMBER: US/10/717,597
; CURRENT FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: US 60/459,782
; PRIOR FILING DATE: 2003-04-03
; PRIOR APPLICATION NUMBER: US 60/427,982
; PRIOR FILING DATE: 2002-11-21
; NUMBER OF SEQ ID NOS: 4904
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 125
; LENGTH: 2587
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-717-597-125

Query Match 76.8%; Score 19.2; DB 7; Length 2587;
Best Local Similarity 87.5%; Pred. No. 56;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAAATCCTGTAGGTTTCAGT 24
|||||
DB 2463 AAAAAAATCCTGTAGGTTTCATT 2440

RESULT 20
US-10-956-157-776/c
; Sequence 776, Application US/10956157
; Publication No. US20050118625A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William
; TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
; TITLE OF INVENTION: HUMAN OSTEOARTHRITIS AND HUMAN PROTEASES
; FILE REFERENCE: 031896-043000 (AM 101081)
; CURRENT APPLICATION NUMBER: US/10/956,157
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 319805
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 776
; LENGTH: 2587
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-956-157-776

Query Match 76.8%; Score 19.2; DB 9; Length 2587;
Best Local Similarity 87.5%; Pred. No. 56;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAAATCCTGTAGGTTTCAGT 24
|||||
DB 2463 AAAAAAATCCTGTAGGTTTCATT 2440

RESULT 21
US-10-981-277-8/c
; Sequence 8, Application US/10981277
; Publication No. US20050181389A1
; GENERAL INFORMATION:
; APPLICANT: Harris, Cole
; APPLICANT: Davis, Lisa
; TITLE OF INVENTION: Compositions and Methods for Glioma Classification
; FILE REFERENCE: 03-968-US
; CURRENT APPLICATION NUMBER: US/10/981,277
; CURRENT FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/516,817
; PRIOR FILING DATE: 2003-11-03
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 8
; LENGTH: 2587
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-981-277-8

Query Match 76.8%; Score 19.2; DB 9; Length 2587;
Best Local Similarity 87.5%; Pred. No. 56;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAAATCCTGTAGGTTTCAGT 24
|||||
DB 2463 AAAAAAATCCTGTAGGTTTCATT 2440

RESULT 22
US-10-252-157-57/c
; Sequence 57, Application US/10252157
; Publication No. US20030190640A1
; GENERAL INFORMATION:
; APPLICANT: Faris, Mary
; APPLICANT: Pearson, Cecelia I.
; TITLE OF INVENTION: GENES EXPRESSED IN PROSTATE CANCER


```

; FILE REFERENCE: PA-0027-1 US
; CURRENT APPLICATION NUMBER: US/10/252,157
; CURRENT FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: 60/295,048
; PRIOR FILING DATE: 2001-05-31
; NUMBER OF SEQ ID NOS: 501
; SOFTWARE: PERL Program
; SEQ ID NO 57
; LENGTH: 3886
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030190640A1 314831.5
;
; NAME/KEY: unsure
; LOCATION: 226-288, 3667-3774
; OTHER INFORMATION: a, t, c, g, or other
US-10-252-157-57

Query Match          76.8%; Score 19.2; DB 6; Length 3886;
Best Local Similarity 87.5%; Pred. No. 59;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAACTCTGTAGGCTTTTCAGT 24
Db 2445 AAAAAATTCCTGTAGGTTTTCATT 2422

RESULT 23
US-09-814-353-20129/c
; Sequence 20129, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20129
; LENGTH: 5569
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1, 2, 3, 4, 5, 6, 7, 5569
; OTHER INFORMATION: n = A,T,C or G
US-09-814-353-20129

Query Match          76.8%; Score 19.2; DB 3; Length 5569;
Best Local Similarity 87.5%; Pred. No. 62;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAACTCTGTAGGCTTTTCAGT 24

```

```

Db 317 AAAAAATTCCTGTAGGTTTTCATT 294

RESULT 24
US-10-357-930-24829/c
; Sequence 24829, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endegre, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24829
; LENGTH: 5570
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1, 2, 3, 4, 5, 6, 7, 5570
; OTHER INFORMATION: n = A,T,C or G
US-10-357-930-24829

Query Match          76.8%; Score 19.2; DB 8; Length 5570;
Best Local Similarity 87.5%; Pred. No. 62;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAACTCTGTAGGCTTTTCAGT 24
Db 317 AAAAAATTCCTGTAGGTTTTCATT 294

RESULT 25
US-10-981-277-36
; Sequence 36, Application US/10981277
; Publication No. US20050181389A1
; GENERAL INFORMATION:
; APPLICANT: Harris, Cole
; APPLICANT: Davis, Lisa
; TITLE OF INVENTION: Compositions and Methods for Glioma Classification
; FILE REFERENCE: 03-968-US
; CURRENT APPLICATION NUMBER: US/10/981,277
; CURRENT FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/516,817
; PRIOR FILING DATE: 2003-11-03
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 36
; LENGTH: 150525
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-981-277-36

```

Query Match 76.8%; Score 19.2; DB 9; Length 150525;
Best Local Similarity 87.5%; Pred. No. 99;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAACTCCTAGGCTTTTCAGT 24
||||| ||||| ||||| ||||| |||||
Db 92490 AAAAAATCCTAGGCTTTTCATT 92513

RESULT 26
US-10-981-277-35
; Sequence 35, Application US/10981277
; Publication No. US20050181389A1
; GENERAL INFORMATION:
; APPLICANT: Harris, Cole
; APPLICANT: Davis, Lisa
; TITLE OF INVENTION: Compositions and Methods for Glioma Classification
; FILE REFERENCE: 03-968-US
; CURRENT APPLICATION NUMBER: US/10/981,277
; CURRENT FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/516,817
; PRIOR FILING DATE: 2003-11-03
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 35
; LENGTH: 166536
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-981-277-35

Query Match 76.8%; Score 19.2; DB 9; Length 166536;
Best Local Similarity 87.5%; Pred. No. 1e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAACTCCTAGGCTTTTCAGT 24
||||| ||||| ||||| ||||| |||||
Db 26151 AAAAAATCCTAGGCTTTTCATT 26174

RESULT 27
US-10-027-632-129877
; Sequence 129877, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 129877
; LENGTH: 594
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-129877

Query Match 74.4%; Score 18.6; DB 5; Length 594;

Best Local Similarity 84.0%; Pred. No. 89;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAACTCCTAGGCTTTTCAGTG 25
||||| ||||| ||||| ||||| |||||
Db 346 AAAAAATCCTCTGGGCTTTCTGTG 370

RESULT 28
US-10-027-632-129878
; Sequence 129878, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 129878
; LENGTH: 594
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-129878

Query Match 74.4%; Score 18.6; DB 5; Length 594;
Best Local Similarity 84.0%; Pred. No. 89;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAACTCCTAGGCTTTTCAGTG 25
||||| ||||| ||||| ||||| |||||
Db 346 AAAAAATCCTCTGGGCTTTCTGTG 370

RESULT 29
US-10-027-632-129877
; Sequence 129877, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28

```

; PRIOR APPLICATION NUMBER: US 60/146,002
;
; PRIORITY FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 129877
; LENGTH: 594
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-129877

```

```
Query Match      74.4%; Score 18.6; DB 6; Length 594;
Best Local Similarity 84.0%; Pred.No. 89;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
```

Qy 1 AAAAACTCCTGTAGGCTTTCAGTG 25
||| ||| ||| ||| ||| ||| ||| |||
Dp 346 AAAAACTCCTGTAGGCTTTCAGTG 370

```

RESULT 30
US-10-027-632-129878
; Sequence 129878, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; ; Sequence 129878, Application US/10027632
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 129878
; LENGTH: 594
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-129878

```

Query Match	74.4%;	Score 18.6;	DB 6;	Length 594;
Best Local Similarity	84.0%;	Pred. No. 89;		
Matches	21;	Conservative	0;	Mismatches 4;
			Indels	0;
			Gaps	0;

Qy 1 AAAAACTCCTGTAGGCTTTTCAGTG 25
||| ||| ||| ||| ||| ||| ||| |||
pb 346 AAAAATCTCCTCTGGGCTTCTGTG 370

```

RESULT 31
US-10-027-632-133274
; Sequence 133274, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108927.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006

```

```

? PRIOR FILING DATE: 2000-07-12
? PRIOR APPLICATION NUMBER: US 60/198,676
? PRIOR FILING DATE: 2000-04-20
? PRIOR APPLICATION NUMBER: US 60/193,483
? PRIOR FILING DATE: 2000-03-29
? PRIOR APPLICATION NUMBER: US 60/185,218
? PRIOR FILING DATE: 2000-02-24
? PRIOR APPLICATION NUMBER: US 60/167,363
? PRIOR FILING DATE: 1999-11-23
? PRIOR APPLICATION NUMBER: US 60/156,358
? PRIOR FILING DATE: 1999-09-28
? PRIOR APPLICATION NUMBER: US 60/146,002
? PRIOR FILING DATE: 1999-08-09
? NUMBER OF SEQ ID NOS: 325720
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 133274
? LENGTH: 823
? TYPE: DNA
? ORGANISM: Human
? US-10-027-632-133274

```

Query Match	74.4%;	Score 18.6;	DB 5;	Length 823;
Best Local Similarity	84.0%;	Pred. No. 94;		
Matches 21: Conservative	0;	Mismatches 4;	Indels 0;	Gaps 0;

QY

1 AAAAACTCCTGTAGGCTTTTCAGTG 25
| | | | | | | | | | | | | | |
pB

460 AGAAACTGCTGGAGGTTTTTCAGTG 484

```

RESULT 32
US-10-027-632-133275
; Sequence 133275, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 133275
; LENGTH: 823
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-133275

```

Query Match	74.4%;	Score 18.6;	DB 5;	Length 823;
Best Local Similarity	84.0%;	Pred. No. 94;		
Matches 21:	Conservative	0: Mismatches	4: Indels	0: Gaps

QY 1 AAAAACTCCTGTAGGCTTTCAGTG 25
| | | | | | | | | |
nB 460 AGAAAACTGCTGGAGGTTTTTCAGTG 48

RESULT 33

```
US-10-027-632-133274
; Sequence 133274, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; POLYMORPHISMS IN THE HUMAN GENOME
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 133274
; LENGTH: 823
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-133274

Query Match      74.4%; Score 18.6; DB 6; Length 823;
Best Local Similarity 84.0%; Pred. No. 94;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAACTCCTGTAGGCTTTTCAGTG 25
   ||||| ||||| ||||| ||||| |||||
Db 460 AGAAAACTGCTGGAGTTTTCAGTG 484

RESULT 35
US-10-369-626-3
; Sequence 3, Application US/10369626
; Publication No. US20030143624A1
; GENERAL INFORMATION:
; APPLICANT: BEASLEY, Ellen M. et al
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
; ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
; THEREOF
; FILE REFERENCE: CL001213DIV
; CURRENT APPLICATION NUMBER: US/10/369,626
; CURRENT FILING DATE: 2003-02-21
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 39982
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-369-626-3

Query Match      74.4%; Score 18.6; DB 6; Length 39982;
Best Local Similarity 84.0%; Pred. No. 1.6e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAACTCCTGTAGGCTTTTCAGTG 25
   ||||| ||||| ||||| ||||| |||||
Db 14081 AAAGAAATCCTGTGGCTTTCTGTG 14105

RESULT 36
US-10-900-420-3
; Sequence 3, Application US/10900420
; Publication No. US2005003434A1
; GENERAL INFORMATION:
; APPLICANT: BEASLEY, Ellen M. et al
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
; ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
; THEREOF
; FILE REFERENCE: CL001213DIV-II
; CURRENT APPLICATION NUMBER: US/10/900,420
; CURRENT FILING DATE: 2004-07-28
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 39982
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-900-420-3

Query Match      74.4%; Score 18.6; DB 8; Length 39982;
Best Local Similarity 84.0%; Pred. No. 1.6e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAACTCCTGTAGGCTTTTCAGTG 25
   ||||| ||||| ||||| ||||| |||||
Db 14081 AAAGAAATCCTGTGGCTTTCTGTG 14105

RESULT 37
US-10-425-115-172275/c
; Sequence 172275, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
```

```
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 172275
; LENGTH: 317
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_88696C.1
US-10-425-115-172275

Query Match          72.8%; Score 18.2; DB 8; Length 317;
Best Local Similarity 87.0%; Pred. No. 1.3e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAACTCTGTAGGCTTTCAG 23
   ||||| ||||| ||||| |||||
Db 38 AAAAACTGCTGTAGGCTGTGAG 16

RESULT 38
US-10-425-115-33082/c
; Sequence 33082, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 33082
; LENGTH: 351
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_130180C.1
US-10-425-115-33082

Query Match          72.8%; Score 18.2; DB 8; Length 351;
Best Local Similarity 87.0%; Pred. No. 1.3e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAACTCTGTAGGCTTTCAG 23
   ||||| ||||| ||||| |||||
Db 102 AAAAACTCTGTAGGCTGTGAG 80

RESULT 39
US-09-925-065A-814695
; Sequence 814695, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
```

```
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 814695
; LENGTH: 429
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-814695

Query Match          72.8%; Score 18.2; DB 4; Length 429;
Best Local Similarity 87.0%; Pred. No. 1.3e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 AAAAATCTCTGTAGGCTTTCAGTG 25
   ||||| ||||| ||||| |||||
Db 22 AAAAATCTCTAAAGGCTATCAGTG 44

RESULT 40
US-10-972-079-59759
; Sequence 59759, Application US/10972079
; Publication No. US20050153317A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: ROSENFELD, David
; APPLICANT: KERR, Richard
; APPLICANT: BATES, Stephen
; APPLICANT: HOLM, Tom
; TITLE OF INVENTION: METHODS & SYSTEMS FOR INFERRING TRAITS TO BREED & MANAGE NON-BEE
; FILE REFERENCE: MM1110-2
; CURRENT APPLICATION NUMBER: US/10/972,079
; CURRENT FILING DATE: 2004-10-22
; PRIOR APPLICATION NUMBER: US 60/514,333
; PRIOR FILING DATE: 2003-10-24
; NUMBER OF SEQ ID NOS: 96631
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 59759
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Chicken 19866894318382_1
US-10-972-079-59759

Query Match          72.8%; Score 18.2; DB 9; Length 600;
Best Local Similarity 87.0%; Pred. No. 1.4e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 AAAAATCTCTGTAGGCTTTCAGTG 25
   ||||| ||||| ||||| |||||
Db 362 AAAAATCTCTGTAGGCTTTCATG 384

RESULT 41
US-09-925-065A-450957
; Sequence 450957, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
```

```
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 450957
; LENGTH: 638
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-450957
```

```
Query Match 72.8%; Score 18.2; DB 4; Length 638;
Best Local Similarity 87.0%; Pred. No. 1.4e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 1 AAAAATCCTCTAGGCTTTTCAG 23
    ||||| ||||| ||||| |||||
Db 348 AAACAAATCCTATAGGCTTTTCAG 370
```

```
RESULT 42
US-09-925-065A-554692
; Sequence 554692, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 554692
; LENGTH: 660
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-554692
```

```
Query Match 72.8%; Score 18.2; DB 4; Length 660;
Best Local Similarity 87.0%; Pred. No. 1.4e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 1 AAAAATCCTCTAGGCTTTTCAG 23
    ||||| ||||| ||||| |||||
Db 264 AAACAAATCCTATAGGCTTTTCAG 286
```

```
RESULT 43
US-09-925-065A-681797
; Sequence 681797, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
```

```
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 681797
; LENGTH: 834
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-681797
```

```
Query Match 72.8%; Score 18.2; DB 4; Length 834;
Best Local Similarity 87.0%; Pred. No. 1.5e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 3 AAAATCCTCTAGGCTTTTCAGTG 25
    ||||| ||||| ||||| |||||
Db 149 AAAATCCTCTAAAGGCTATCAGTG 171
```

```
RESULT 44
US-10-027-632-121229
; Sequence 121229, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 121229
; LENGTH: 919
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-121229
```

```
Query Match 72.8%; Score 18.2; DB 5; Length 919;
Best Local Similarity 87.0%; Pred. No. 1.5e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 3 AAAATCCTCTAGGCTTTTCAGTG 25
    ||||| ||||| ||||| |||||
Db 645 AAAGAGCTGTAGGCTTTTCAGTG 667
```

```
RESULT 45
US-10-027-632-121230
; Sequence 121230, Application US/10027632
; Publication No. US20020198371A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 121230
; LENGTH: 919
; TYPE: DNA
; ORGANISM: Human
; US-10-027-632-121230

Query Match          72.8%; Score 18.2; DB 5; Length 919;
Best Local Similarity 87.0%; Pred. No. 1.5e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 AAAAATCCTGTAGGCTTTCAGTG 25
   ||||| ||||| ||||| |||||
Db 645 AAAAGAGCTGTAGGCTTTCAGTG 667

RESULT 46
US-10-027-632-121229
; Sequence 121229, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 121229
; LENGTH: 919
; TYPE: DNA
; ORGANISM: Human
; US-10-027-632-121229

Query Match          72.8%; Score 18.2; DB 6; Length 919;
Best Local Similarity 87.0%; Pred. No. 1.5e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 AAAAATCCTGTAGGCTTTCAGTG 25
   ||||| ||||| ||||| |||||
Db 645 AAAAGAGCTGTAGGCTTTCAGTG 667

RESULT 47
US-10-027-632-121230
; Sequence 121230, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 121230
; LENGTH: 919
; TYPE: DNA
; ORGANISM: Human
; US-10-027-632-121230

Query Match          72.8%; Score 18.2; DB 6; Length 919;
Best Local Similarity 87.0%; Pred. No. 1.5e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 AAAAATCCTGTAGGCTTTCAGTG 25
   ||||| ||||| ||||| |||||
Db 645 AAAAGAGCTGTAGGCTTTCAGTG 667

RESULT 48
US-10-027-632-117480/C
; Sequence 117480, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
```

```
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 121230
; LENGTH: 919
; TYPE: DNA
; ORGANISM: Human
; US-10-027-632-121230

Query Match          72.8%; Score 18.2; DB 6; Length 919;
Best Local Similarity 87.0%; Pred. No. 1.5e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 AAAAATCCTGTAGGCTTTCAGTG 25
   ||||| ||||| ||||| |||||
Db 645 AAAAGAGCTGTAGGCTTTCAGTG 667

RESULT 47
US-10-027-632-117480/C
; Sequence 117480, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
```

```
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 117480
; LENGTH: 1136
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1136)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-117480
```

```
Query Match          72.8%; Score 18.2; DB 5; Length 1136;
Best Local Similarity 87.0%; Pred. No. 1.5e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
QY      3 AAAACTCCTGTAGGCTTTTCAGTG 25
          ||||| ||||| ||||| ||||| |||||
Db      573 AAAACTCCTTTATGTTTTCAGTG 551
```

RESULT 49

```
US-10-027-632-117480/c
; Sequence 117480, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 117480
; LENGTH: 1136
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1136)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-117480
```

```
Query Match          72.8%; Score 18.2; DB 6; Length 1136;
Best Local Similarity 87.0%; Pred. No. 1.5e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
QY      3 AAAACTCCTGTAGGCTTTTCAGTG 25
          ||||| ||||| ||||| ||||| |||||
Db      573 AAAACTCCTTTATGTTTTCAGTG 551
```

RESULT 50

```
US-10-424-599-42397/c
; Sequence 42397, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
```

```
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 42397
; LENGTH: 1175
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_138285C.1
US-10-424-599-42397
```

```
Query Match          72.8%; Score 18.2; DB 7; Length 1175;
Best Local Similarity 87.0%; Pred. No. 1.6e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
QY      2 AAAAATCCTCTGTAGGCTTTTCAGT 24
          ||||| ||||| ||||| ||||| |||||
Db      587 AAAACACCTATAGGCTTTTCAGT 565
```

Search completed: February 3, 2006, 15:43:52
Job time : 371.556 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 3, 2006, 11:35:31 ; Search time 343.111 Seconds
(without alignments)
61.064 Million cell updates/sec

Title: US-10-719-900-20

Perfect score: 25

Sequence: 1 aaaaactcctgtaggcttcagtg 25

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 6068529 seqs, 419036697 residues

Total number of hits satisfying chosen parameters: 12137058

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database : Published Applications NA New.*

1: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
2: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	18.6	74.4	4499	7	US-10-750-185-43501
C 2	18.6	74.4	4499	7	US-10-750-623-43501
C 3	17.6	70.4	568	8	US-11-128-061-1900
C 4	17.6	70.4	568	8	US-11-128-061-5542
C 5	17.6	70.4	568	8	US-11-128-049-1900
C 6	17.6	70.4	568	8	US-11-128-049-5542
C 7	17.2	68.8	6138	8	US-11-180-000-3
C 8	17.2	68.8	6141	8	US-11-180-000-3
C 9	17.2	68.8	7646	8	US-11-180-000-1
C 10	17.2	68.8	16570	8	US-11-180-000-2
C 11	17.2	68.8	68123	7	US-10-995-561-13348
C 12	17	68.0	201	8	US-11-124-368A-19365
C 13	17	68.0	201	8	US-11-124-368A-19366
C 14	17	68.0	1350	7	US-10-750-185-52644
C 15	17	68.0	1350	7	US-10-750-623-52644
C 16	17	68.0	1516	7	US-10-750-185-39342
C 17	17	68.0	1516	7	US-10-750-623-39342
C 18	17	68.0	1570	7	US-10-750-185-36386
C 19	17	68.0	1570	7	US-10-750-623-36386
C 20	17	68.0	1728	7	US-10-750-185-43815
C 21	17	68.0	1728	7	US-10-750-623-43815
C 22	17	68.0	1954	7	US-10-750-185-43339

Sequence 43339, A	7	US-10-750-623-43339	1954	68.0	17
Sequence 5, Appli	8	US-11-117-169-5	2064	68.0	17
Sequence 7, Appli	8	US-11-117-169-7	2186	68.0	17
Sequence 51243, A	7	US-10-750-185-51243	2417	68.0	17
Sequence 51243, A	7	US-10-750-623-51243	2417	68.0	17
Sequence 486, App	8	US-11-052-554A-486	2526	68.0	17
Sequence 64396, A	7	US-10-750-185-64398	2843	68.0	17
Sequence 64398, A	7	US-10-750-623-64398	2843	68.0	17
Sequence 14, Appli	8	US-11-117-169-14	3068	68.0	17
Sequence 393885,	25	US-11-121-849-393885	367.2	67.2	16.8
Sequence 52818, A	7	US-10-750-185-52818	747	67.2	16.8
Sequence 52818, A	7	US-10-750-623-52818	747	67.2	16.8
Sequence 4713, Ap	7	US-10-750-185-4713	600	66.4	16.6
Sequence 2131, A	7	US-10-750-185-2131	600	66.4	16.6
Sequence 4713, Ap	7	US-10-750-623-4713	600	66.4	16.6
Sequence 2131, A	7	US-10-750-623-2131	600	66.4	16.6
Sequence 26142, A	7	US-10-750-185-26142	781	66.4	16.6
Sequence 26142, A	7	US-10-750-623-26142	781	66.4	16.6
Sequence 52730, A	7	US-10-750-185-52730	914	66.4	16.6
Sequence 52730, A	7	US-10-750-623-52730	914	66.4	16.6
Sequence 57290, A	7	US-10-750-185-57290	1153	66.4	16.6
Sequence 57290, A	7	US-10-750-623-57290	1153	66.4	16.6
Sequence 44378, A	7	US-10-750-185-44378	1496	66.4	16.6
Sequence 44378, A	7	US-10-750-623-44378	1496	66.4	16.6
Sequence 52605, A	7	US-10-750-185-52605	1507	66.4	16.6
Sequence 52605, A	7	US-10-750-623-52605	1507	66.4	16.6
Sequence 50030, A	7	US-10-750-185-50030	1695	66.4	16.6
Sequence 50030, A	7	US-10-750-623-50030	1695	66.4	16.6
Sequence 42351, A	7	US-10-750-185-42351	2251	66.4	16.6
Sequence 42351, A	7	US-10-750-623-42351	2251	66.4	16.6
Sequence 50907, A	7	US-10-750-185-50907	4168	66.4	16.6
Sequence 50907, A	7	US-10-750-623-50907	4168	66.4	16.6
Sequence 13366, A	7	US-10-995-561-13366	83712	66.4	16.6
Sequence 17, Appli	8	US-11-121-086-17	178877	66.4	16.6
Sequence 1, Appli	8	US-11-120-925-1	188056	66.4	16.6
Sequence 34140, A	7	US-10-750-185-34140	969	64.8	16.2
Sequence 34140, A	7	US-10-750-623-34140	969	64.8	16.2
Sequence 53981, A	7	US-10-750-185-53981	1028	64.8	16.2
Sequence 53981, A	7	US-10-750-623-53981	1028	64.8	16.2
Sequence 25195, A	7	US-10-750-185-25195	1032	64.8	16.2
Sequence 25195, A	7	US-10-750-623-25195	1032	64.8	16.2
Sequence 52953, A	7	US-10-750-185-52953	1324	64.8	16.2
Sequence 52953, A	7	US-10-750-623-52953	1324	64.8	16.2
Sequence 52504, A	7	US-10-750-185-52504	1881	64.8	16.2
Sequence 52504, A	7	US-10-750-623-52504	1881	64.8	16.2
Sequence 35931, A	7	US-10-750-185-35931	2303	64.8	16.2
Sequence 35931, A	7	US-10-750-623-35931	2303	64.8	16.2
Sequence 31466, A	7	US-10-750-185-31466	2656	64.8	16.2
Sequence 31466, A	7	US-10-750-623-31466	2656	64.8	16.2
Sequence 576, App	8	US-11-136-527-576	2881	64.8	16.2
Sequence 3, Appli	8	US-11-136-527-576	2881	64.8	16.2
Sequence 1, Appli	8	US-11-176-253-1	17004	64.8	16.2
Sequence 29, Appli	8	US-11-176-253-1	17004	64.8	16.2
Sequence 3271, Ap	8	US-11-091-018-1	611587	64.8	16.2
Sequence 58951, A	7	US-10-793-626-3271	1691140	64.8	16.2
Sequence 58951, A	7	US-10-750-185-58951	567	64.0	16.0
Sequence 48397, A	7	US-10-750-623-48397	884	64.0	16.0
Sequence 48397, A	7	US-10-750-185-48397	884	64.0	16.0
Sequence 40232, A	7	US-10-750-623-40232	1028	64.0	16.0
Sequence 40232, A	7	US-10-750-185-40232	1028	64.0	16.0
Sequence 63598, A	7	US-10-750-623-63598	1141	64.0	16.0
Sequence 63598, A	7	US-10-750-185-63598	1141	64.0	16.0
Sequence 9, Appli	7	US-10-467-657-9	1227	64.0	16.0
Sequence 5467, Ap	7	US-10-467-657-5467	1227	64.0	16.0
Sequence 51455, A	7	US-10-750-185-51455	1245	64.0	16.0
Sequence 51455, A	7	US-10-750-623-51455	1245	64.0	16.0
Sequence 28818, A	7	US-10-750-185-28818	1338	64.0	16.0
Sequence 28818, A	7	US-10-750-623-28818	1338	64.0	16.0
Sequence 26, Appli	7	US-10-978-3608-26	1340	64.0	16.0
Sequence 373, Appl	5	US-10-131-826A-373	1448	64.0	16.0
Sequence 181, App	7	US-10-131-826A-181	1572	64.0	16.0
Sequence 56730, A	7	US-10-750-185-56730	1630	64.0	16.0

; PRIOR FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: EP 10116044.7
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: EP 01104171.2
; PRIOR FILING DATE: 2001-02-21
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 6138
; TYPE: DNA
; ORGANISM: Human cytomegalovirus
US-11-180-000-4

Query Match 68.8%; Score 17.2; DB 8; Length 6138;
Best Local Similarity 86.4%; Pred. No. 71;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAACTCTGTAGGCTTTCA 22
||||| ||||| ||||| |||||
DB 1781 AAAAACTCTGTGGCGTTCA 1760

RESULT 8

US-11-180-000-3/c
; Sequence 3, Application US/11180000
; Publication No. US20060019369A1
; GENERAL INFORMATION:
; APPLICANT: Hahn, Gabriele
; TITLE OF INVENTION: Recombinant vector containing infectious human cytomegalovirus ge
; TITLE OF INVENTION: With preserved wild-type characteristics of clinical isolates
; FILE REFERENCE: 2923-212
; CURRENT APPLICATION NUMBER: US/11/180,000
; CURRENT FILING DATE: 2005-07-13
; PRIOR APPLICATION NUMBER: US/10/275,287
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: PCT/EP02/01867
; PRIOR FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: EP 10116044.7
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: EP 01104171.2
; PRIOR FILING DATE: 2001-02-21
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 6141
; TYPE: DNA
; ORGANISM: Human cytomegalovirus
US-11-180-000-3

Query Match 68.8%; Score 17.2; DB 8; Length 6141;
Best Local Similarity 86.4%; Pred. No. 71;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAACTCTGTAGGCTTTCA 22
||||| ||||| ||||| |||||
DB 1781 AAAAACTCTGTGGCGTTCA 1760

RESULT 9

US-11-180-000-1/c
; Sequence 1, Application US/11180000
; Publication No. US20060019369A1
; GENERAL INFORMATION:
; APPLICANT: Hahn, Gabriele
; TITLE OF INVENTION: Recombinant vector containing infectious human cytomegalovirus ge
; TITLE OF INVENTION: With preserved wild-type characteristics of clinical isolates
; FILE REFERENCE: 2923-212
; CURRENT APPLICATION NUMBER: US/11/180,000
; CURRENT FILING DATE: 2005-07-13
; PRIOR APPLICATION NUMBER: US/10/275,287
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: PCT/EP02/01867
; PRIOR FILING DATE: 2002-02-21

; PRIOR APPLICATION NUMBER: EP 10116044.7
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: EP 01104171.2
; PRIOR FILING DATE: 2001-02-21
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 7646
; TYPE: DNA
; ORGANISM: Human cytomegalovirus
US-11-180-000-1

Query Match 68.8%; Score 17.2; DB 8; Length 7646;
Best Local Similarity 86.4%; Pred. No. 75;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAACTCTGTAGGCTTTCA 22
||||| ||||| ||||| |||||
DB 3273 AAAAACTCTGTGGCGTTCA 3252

RESULT 10

US-11-180-000-2/c
; Sequence 2, Application US/11180000
; Publication No. US20060019369A1
; GENERAL INFORMATION:
; APPLICANT: Hahn, Gabriele
; TITLE OF INVENTION: Recombinant vector containing infectious human cytomegalovirus g
; TITLE OF INVENTION: With preserved wild-type characteristics of clinical isolates
; FILE REFERENCE: 2923-212
; CURRENT APPLICATION NUMBER: US/11/180,000
; CURRENT FILING DATE: 2005-07-13
; PRIOR APPLICATION NUMBER: US/10/275,287
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: PCT/EP02/01867
; PRIOR FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: EP 10116044.7
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: EP 01104171.2
; PRIOR FILING DATE: 2001-02-21
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 16570
; TYPE: DNA
; ORGANISM: Human cytomegalovirus
US-11-180-000-2

Query Match 68.8%; Score 17.2; DB 8; Length 16570;
Best Local Similarity 86.4%; Pred. No. 88;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAACTCTGTAGGCTTTCA 22
||||| ||||| ||||| |||||
DB 12206 AAAAACTCTGTGGCGTTCA 12185

RESULT 11

US-10-995-561-13348
; Sequence 13348, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13348
; LENGTH: 68123

					Mismatches	5;	Indels	0;	Gaps	0;
<hr/>										
QY	1	AAAAAACTCCTGTAGGCTTTCA	TG	25						
DB	162	AATAATCTCATGTATGTATTTCAGT	G	186						
<hr/>										
RESULT 14										
US-10-750-185-52644										
; Sequence 52644, Application US/10750185										
; Publication No. US20050260603A1										
; GENERAL INFORMATION:										
; APPLICANT: MMI GENOMICS, INC.										
; APPLICANT: DENISE, Sue K.										
; APPLICANT: KERR, Richard										
; APPLICANT: ROSENFELD, David										
; APPLICANT: HOLM, Tom										
; APPLICANT: BATES, Stephen										
; APPLICANT: FANTIN, Dennis										
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS										
; FILE REFERENCE: MMII100-2										
; CURRENT APPLICATION NUMBER: US/10750,185										
; CURRENT FILING DATE: 2003-12-31										
; PRIOR APPLICATION NUMBER: US 60/437,482										
; PRIOR FILING DATE: 2002-12-31										
; NUMBER OF SEQ ID NOS: 64922										
; SOFTWARE: PatentIn version 3.1										
; SEQ ID NO 52644										
; LENGTH: 1350										
; TYPE: DNA										
; ORGANISM: Bovine 19866880826885										
US-10-750-185-52644										
<hr/>										
Query Match 68.0%; Score 17; DB 7; Length 1350;										
Best Local Similarity 80.0%; Pred. No. 66;										
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;										
<hr/>										
QY	1	AAAAAACTCCTGTAGGCTTTCA	TG	25						
DB	912	AGAAAACCTCTCTGGCTTTCCAGAG	936							
<hr/>										
RESULT 15										
US-10-750-623-52644										
; Sequence 52644, Application US/10750623										
; Publication No. US20050287531A1										
; GENERAL INFORMATION:										
; APPLICANT: MMI GENOMICS, INC.										
; APPLICANT: DENISE, Sue K.										
; APPLICANT: KERR, Richard										
; APPLICANT: ROSENFELD, David										
; APPLICANT: HOLM, Tom										
; APPLICANT: BATES, Stephen										
; APPLICANT: FANTIN, Dennis										
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS										
; FILE REFERENCE: MMII100-1										
; CURRENT APPLICATION NUMBER: US/10750,623										
; CURRENT FILING DATE: 2003-12-31										
; PRIOR APPLICATION NUMBER: US 60/437,482										
; PRIOR FILING DATE: 2002-12-31										
; NUMBER OF SEQ ID NOS: 64922										
; SOFTWARE: PatentIn version 3.1										
; SEQ ID NO 52644										
; LENGTH: 1350										
; TYPE: DNA										
; ORGANISM: Bovine 19866880826885										
US-10-750-623-52644										
<hr/>										
Query Match 68.0%; Score 17; DB 7; Length 1350;										
Best Local Similarity 80.0%; Pred. No. 66;										
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;										

US-10-750-185-43815
; Sequence 43815, Application US/10750185
; Publication No. US20050260603A1

GENERAL INFORMATION:

; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis

; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS

; FILE REFERENCE: MM1100-2

; CURRENT APPLICATION NUMBER: US/10/750,185

; CURRENT FILING DATE: 2003-12-31

; PRIOR APPLICATION NUMBER: US 60/437,482

; PRIOR FILING DATE: 2002-12-31

; NUMBER OF SEQ ID NOS: 64922

; SOFTWARE: Patentin version 3.1

; SEQ ID NO 43815

; LENGTH: 1728

; TYPE: DNA

; ORGANISM: Bovine 19866880856399

US-10-750-185-43815

Query Match 68.0%; Score 17; DB 7; Length 1728;

Best Local Similarity 80.0%; Pred. No. 69;

Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAACTCTGTAGGCTTTTCAGTG 25

Db 44 AAAAAATTCAGTATCCTTTTCAGTG 68

RESULT 21

US-10-750-623-43815

; Sequence 43815, Application US/10750623

; Publication No. US20050287531A1

GENERAL INFORMATION:

; APPLICANT: MMI GENOMICS, INC.

; APPLICANT: DENISE, Sue K.

; APPLICANT: KERR, Richard

; APPLICANT: ROSENFELD, David

; APPLICANT: HOLM, Tom

; APPLICANT: BATES, Stephen

; APPLICANT: FANTIN, Dennis

; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS

; FILE REFERENCE: MM1100-1

; CURRENT APPLICATION NUMBER: US/10/750,623

; CURRENT FILING DATE: 2003-12-31

; PRIOR APPLICATION NUMBER: US 60/437,482

; PRIOR FILING DATE: 2002-12-31

; NUMBER OF SEQ ID NOS: 64922

; SOFTWARE: Patentin version 3.1

; SEQ ID NO 43815

; LENGTH: 1728

; TYPE: DNA

; ORGANISM: Bovine 19866880856399

US-10-750-623-43815

Query Match 68.0%; Score 17; DB 7; Length 1728;

Best Local Similarity 80.0%; Pred. No. 69;

Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAACTCTGTAGGCTTTTCAGTG 25

Db 44 AAAAAATTCAGTATCCTTTTCAGTG 68

RESULT 22

US-10-750-185-43339/c

; Sequence 43339, Application US/10750185

; Publication No. US20050260603A1

GENERAL INFORMATION:

; APPLICANT: MMI GENOMICS, INC.

; APPLICANT: DENISE, Sue K.

; APPLICANT: KERR, Richard

; APPLICANT: ROSENFELD, David

; APPLICANT: HOLM, Tom

; APPLICANT: BATES, Stephen

; APPLICANT: FANTIN, Dennis

; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS

; FILE REFERENCE: MM1100-2

; CURRENT APPLICATION NUMBER: US/10/750,185

; CURRENT FILING DATE: 2003-12-31

; PRIOR APPLICATION NUMBER: US 60/437,482

; PRIOR FILING DATE: 2002-12-31

; NUMBER OF SEQ ID NOS: 64922

; SOFTWARE: Patentin version 3.1

; SEQ ID NO 43339

; LENGTH: 1954

; TYPE: DNA

; ORGANISM: Bovine 19866880591323

US-10-750-185-43339

Query Match 68.0%; Score 17; DB 7; Length 1954;

Best Local Similarity 80.0%; Pred. No. 71;

Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAACTCTGTAGGCTTTTCAGTG 25

Db 1748 AAAAAATGCTAAAGGATTTTCAGTG 1724

RESULT 23

US-10-750-623-43339/c

; Sequence 43339, Application US/10750623

; Publication No. US20050287531A1

GENERAL INFORMATION:

; APPLICANT: MMI GENOMICS, INC.

; APPLICANT: DENISE, Sue K.

; APPLICANT: KERR, Richard

; APPLICANT: ROSENFELD, David

; APPLICANT: HOLM, Tom

; APPLICANT: BATES, Stephen

; APPLICANT: FANTIN, Dennis

; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS

; FILE REFERENCE: MM1100-1

; CURRENT APPLICATION NUMBER: US/10/750,623

; CURRENT FILING DATE: 2003-12-31

; PRIOR APPLICATION NUMBER: US 60/437,482

; PRIOR FILING DATE: 2002-12-31

; NUMBER OF SEQ ID NOS: 64922

; SOFTWARE: Patentin version 3.1

; SEQ ID NO 43339

; LENGTH: 1954

; TYPE: DNA

; ORGANISM: Bovine 19866880591323

US-10-750-623-43339

Query Match 68.0%; Score 17; DB 7; Length 1954;

Best Local Similarity 80.0%; Pred. No. 71;

Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAACTCTGTAGGCTTTTCAGTG 25

Db 1748 AAAAAATGCTAAAGGATTTTCAGTG 1724

RESULT 24

US-11-117-169-5/c

; Sequence 5, Application US/11117169

; Publication No. US2005026609A1

GENERAL INFORMATION:

; APPLICANT: University of South Florida

; APPLICANT: Mohapatra, Shyam

;/ TITLE OF INVENTION: Nanogene Therapy For Cell Proliferation Disorders
;/ FILE REFERENCE: USF-216PTCZ
;/ CURRENT APPLICATION NUMBER: US/11/117,169
;/ CURRENT FILING DATE: 2005-04-27
;/ PRIOR APPLICATION NUMBER: 60/565,756
;/ PRIOR FILING DATE: 2004-04-27
;/ NUMBER OF SEQ ID NOS: 31
;/ SOFTWARE: PatentIn version 3.3
;/ SEQ ID NO 5
;/ LENGTH: 2064
;/ TYPE: DNA
;/ ORGANISM: Homo sapiens
;/ FEATURE:
;/ NAME/KEY: CDS
;/ LOCATION: (1)..(2064)
;/ PUBLICATION INFORMATION:
;/ AUTHORS: Hovnanian, A., et al.
;/ TITLE: The human 2', 5'-oligoadenylate synthetase locus is composed of
;/ TITLE: three distinct genes
;/ JOURNAL: Genomics
;/ VOLUME: 52
;/ ISSUE: 3
;/ PAGES: 267-277
;/ DATE: 1998
;/ DATABASE ACCESSION NUMBER: NCBI/NM_002535
;/ DATABASE ENTRY DATE: 2003-04-03
;/ PUBLICATION INFORMATION:
;/ AUTHORS: Marie, I. and Hovanessian, A.G.
;/ TITLE: The 69-kDa 2-5A synthetase is composed of two homologous and
;/ TITLE: adjacent functional domains
;/ JOURNAL: J. Biol. Chem.
;/ VOLUME: 267
;/ ISSUE: 14
;/ PAGES: 9933-9939
;/ DATE: 1992
;/ DATABASE ACCESSION NUMBER: NCBI/NM_002535
;/ DATABASE ENTRY DATE: 2003-04-03
;/ PUBLICATION INFORMATION:
;/ AUTHORS: Marie, I., et al.
;/ TITLE: Differential expression and distinct structure of 69- and 100-kDa
;/ TITLE: forms of 2-5A synthetase
;/ JOURNAL: J. Biol. Chem.
;/ VOLUME: 265
;/ ISSUE: 30
;/ PAGES: 18601-18607
;/ DATE: 1990
;/ DATABASE ACCESSION NUMBER: NCBI/NM_002535
;/ DATABASE ENTRY DATE: 2003-04-03
;/ PUBLICATION INFORMATION:
;/ AUTHORS: Marie I., et al.
;/ TITLE: Preparation and characterization of polyclonal antibodies
;/ TITLE: specific for the 69 and 100 k-dalton forms of human 2-5A
;/ TITLE: synthetase
;/ JOURNAL: Biochem. Biophys. Res. Commun.
;/ VOLUME: 160
;/ ISSUE: 2
;/ PAGES: 580-587
;/ DATE: 1989
;/ DATABASE ACCESSION NUMBER: NCBI/NM_002535
;/ DATABASE ENTRY DATE: 2003-04-03
;/ PUBLICATION INFORMATION:
;/ AUTHORS: Hovanessian, A.G., et al.
;/ TITLE: Characterization of 69- and 100-kDa forms of 2-5A-synthetase from
;/ TITLE: interferon-treated human cells
;/ JOURNAL: J. Biol. Chem.
;/ VOLUME: 263
;/ ISSUE: 10
;/ PAGES: 4959
;/ DATE: 1988
;/ DATABASE ACCESSION NUMBER: NCBI/NM_002535
;/ DATABASE ENTRY DATE: 2003-04-03
;/ PUBLICATION INFORMATION:
;/ AUTHORS: Hovanessian, A.G., et al.

;/ TITLE: Identification of 69-kd and 100-kd forms of 2-5A synthetase
;/ JOURNAL: EMBO J.
;/ VOLUME: 6
;/ ISSUE: 5
;/ PAGES: 1273-1280
;/ DATE: 1987
;/ DATABASE ACCESSION NUMBER: NCBI/NM_002535
;/ DATABASE ENTRY DATE: 2003-04-03
;/ US-11-117-169-5

Query Match 68.0%; Score 17; DB 8; Length 2064;
Best Local Similarity 80.0%; Pred.No.72;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAAAATCCTGTAGGCTTTCAGTG 25
|||||
Db 567 AAAAAAATCCTGTAGGCTTTCAGTG 543
|||||

RESULT 25
US-11-117-169-7/c
;/ Sequence 7, Application US/11117169
;/ Publication No. US20050266093A1
;/ GENERAL INFORMATION:
;/ APPLICANT: University of South Florida
;/ APPLICANT: Mohapatra, Shyam
;/ TITLE OF INVENTION: Nanogene Therapy For Cell Proliferation Disorders
;/ FILE REFERENCE: USF-216PTCZ
;/ CURRENT APPLICATION NUMBER: US/11/117,169
;/ CURRENT FILING DATE: 2005-04-27
;/ PRIOR APPLICATION NUMBER: 60/565,756
;/ PRIOR FILING DATE: 2004-04-27
;/ NUMBER OF SEQ ID NOS: 31
;/ SOFTWARE: PatentIn version 3.3
;/ SEQ ID NO 7
;/ LENGTH: 2186
;/ TYPE: DNA
;/ ORGANISM: Homo sapiens
;/ PUBLICATION INFORMATION:
;/ AUTHORS: Marie, I. and Hovanessian, A.G.
;/ TITLE: The 69-kDa 2-5A synthetase is composed of two homologous and
;/ TITLE: adjacent functional domains
;/ JOURNAL: J. Biol. Chem.
;/ VOLUME: 267
;/ ISSUE: 14
;/ PAGES: 9933-9939
;/ DATE: 1992
;/ DATABASE ACCESSION NUMBER: (unknown)
;/ DATABASE ENTRY DATE: 2003-04-03
;/ US-11-117-169-7

Query Match 68.0%; Score 17; DB 8; Length 2186;
Best Local Similarity 80.0%; Pred.No.73;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAAAATCCTGTAGGCTTTCAGTG 25
|||||
Db 567 AAAAAAATCCTGTAGGCTTTCAGTG 543
|||||

RESULT 26
US-10-750-185-51243/c
;/ Sequence 51243, Application US/10750185
;/ Publication No. US20050266093A1
;/ GENERAL INFORMATION:
;/ APPLICANT: MMI GENOMICS, INC.
;/ APPLICANT: DENISE, Sue K.
;/ APPLICANT: KERR, Richard
;/ APPLICANT: ROSENFELD, David
;/ APPLICANT: HOLM, Tom
;/ APPLICANT: BATES, Stephen
;/ APPLICANT: FANTIN, Dennis
;/ TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS


```
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 51243
; LENGTH: 2417
; TYPE: DNA
; ORGANISM: Bovine 19866880878023
US-10-750-185-51243
```

```
Query Match 68.0%; Score 17; DB 7; Length 2417;
Best Local Similarity 80.0%; Pred. No. 74;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
```

```
QY 1 AAAAACTCCTGTAGGCTTTTCAGTG 25
||||| ||||| ||||| ||||| |||||
Db 519 AAAAAACAGGAGTAGGTTTTCAGTG 495
```

```
RESULT 27
US-10-750-623-51243/c
; Sequence 51243, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 51243
; LENGTH: 2417
; TYPE: DNA
; ORGANISM: Bovine 19866880878023
US-10-750-623-51243
```

```
Query Match 68.0%; Score 17; DB 7; Length 2417;
Best Local Similarity 80.0%; Pred. No. 74;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
```

```
QY 1 AAAAACTCCTGTAGGCTTTTCAGTG 25
||||| ||||| ||||| ||||| |||||
Db 519 AAAAAACAGGAGTAGGTTTTCAGTG 495
```

```
RESULT 28
US-11-052-554A-486
; Sequence 486, Application US/11052554A
; Publication No. US2005028866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: PROTEINS OF THERAPEUTIC POTENTIAL
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; CURRENT FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
```

```
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 486
; LENGTH: 2526
; TYPE: DNA
; ORGANISM: Helicobacter pylori J99
US-11-052-554A-486
```

```
Query Match 68.0%; Score 17; DB 8; Length 2526;
Best Local Similarity 80.0%; Pred. No. 75;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
```

```
QY 1 AAAAACTCCTGTAGGCTTTTCAGTG 25
||||| ||||| ||||| ||||| |||||
Db 7 AAAAAAGTTCTGTGGCTTTTAACTG 31
```

```
RESULT 29
US-10-750-185-64398
; Sequence 64398, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 64398
; LENGTH: 2843
; TYPE: DNA
; ORGANISM: Bovine 19866880746758
US-10-750-185-64398
```

```
Query Match 68.0%; Score 17; DB 7; Length 2843;
Best Local Similarity 80.0%; Pred. No. 77;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
```

```
QY 1 AAAAACTCCTGTAGGCTTTTCAGTG 25
||||| ||||| ||||| ||||| |||||
Db 717 AAAAACTCCTTTAAGCATTTCAGAG 741
```

```
RESULT 30
US-10-750-623-64398
; Sequence 64398, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 64398
; LENGTH: 2843
```

```

; TYPE: DNA
; ORGANISM: Bovine 19866880746758
US-10-750-623-64398

Query Match      68.0%   Score 17; DB 7; Length 2843;
Best Local Similarity 80.0%; Pred.No.77;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps

QY      1 AAAAACTCCTGTAGGCTTTTCAGTG 25
      ||||| ||||| ||||| |||||
DB      717 AAAAACTCCTTTAAGCATTCAGAG 741

RESULT 31
US-11-117-169-14/C
; Sequence 14, Application US/11117169
; Publication No. US20050266093A1
; GENERAL INFORMATION:
; APPLICANT: University of South Florida
; APPLICANT: Mohapatra, Shyam
; TITLE OF INVENTION: Nanogene Therapy For Cell Proliferation Disorders
; FILE REFERENCE: USF-216PTCZ
; CURRENT APPLICATION NUMBER: US/11/117,169
; CURRENT FILING DATE: 2005-04-27
; PRIOR APPLICATION NUMBER: 60/565,756
; PRIOR FILING DATE: 2004-04-27
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 14
; LENGTH: 3068
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; AUTHORS: Hovnanian, A., et al.
; TITLE: The human 2',5'-oligoadenylate synthetase locus is composed of
; TITLE: three distinct genes
; JOURNAL: Genomics
; VOLUME: 52
; ISSUE: 3
; PAGES: 267-277
; DATE: 1998
; DATABASE accession number: NCBI/NM_002535
; DATABASE entry date: 2003-04-03
; PUBLICATION INFORMATION:
; AUTHORS: Marie, I. and Hovanessian, A.G.
; TITLE: The 69-kDa 2-5A synthetase is composed of two homologous and
; TITLE: adjacent functional domains
; JOURNAL: J. Biol. Chem.
; VOLUME: 267
; ISSUE: 14
; PAGES: 9933-9939
; DATE: 1992
; DATABASE accession number: NCBI/NM_002535
; DATABASE entry date: 2003-04-03
; PUBLICATION INFORMATION:
; AUTHORS: Marie, I., et al.
; TITLE: Differential expression and distinct structure of 69- and 100-kDa
; TITLE: forms of 2-5A synthetase
; JOURNAL: J. Biol. Chem.
; VOLUME: 265
; ISSUE: 30
; PAGES: 18601-18607
; DATE: 1990
; DATABASE accession number: NCBI/NM_002535
; DATABASE entry date: 2003-04-03
; PUBLICATION INFORMATION:
; AUTHORS: Marie, I., et al.
; TITLE: Preparation and characterization of polyclonal antibodies
; JOURNAL: Biochem. Biophys. Res. Commun.
; VOLUME: 160
; ISSUE: 2
; PAGES: 580-587
; DATE: 1989

```

; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 52818
; LENGTH: 747
; TYPE: DNA
; ORGANISM: Bovine 19866881647514
US-10-750-185-52818

Query Match 67.2%; Score 16.8; DB 7; Length 747;
Best Local Similarity 90.0%; Pred. No. 73;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 AAACCTCTGTAGGCTTTTCAG 23
DB 564 AAACCTCTGTAGTCTTTTCAG 545

RESULT 34
US-10-750-623-52818/c
; Sequence 52818, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 52818
; LENGTH: 747
; TYPE: DNA
; ORGANISM: Bovine 19866881647514
US-10-750-623-52818

Query Match 67.2%; Score 16.8; DB 7; Length 747;
Best Local Similarity 90.0%; Pred. No. 73;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 AAACCTCTGTAGGCTTTTCAG 23
DB 564 AAACCTCTGTAGTCTTTTCAG 545

RESULT 35
US-10-750-185-4713/c
; Sequence 4713, Application US/10750185
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31

; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 4713
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Bovine MMBT04380
US-10-750-185-4713

Query Match 66.4%; Score 16.6; DB 7; Length 600;
Best Local Similarity 82.6%; Pred. No. 88;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 AAAACCTCTGTAGGCTTTTCAGTG 25
DB 129 AATACTCTCTGGGGCTTCCAGTG 107

RESULT 36
US-10-750-185-21331/c
; Sequence 21331, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 21331
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Bovine MMBT04380
US-10-750-185-21331

Query Match 66.4%; Score 16.6; DB 7; Length 600;
Best Local Similarity 82.6%; Pred. No. 88;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 AAAACCTCTGTAGGCTTTTCAGTG 25
DB 129 AATACTCTCTGGGGCTTCCAGTG 107

RESULT 37
US-10-750-623-4713/c
; Sequence 4713, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31

; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4713
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Bovine MMBT04380
US-10-750-623-4713

Query Match 66.4%; Score 16.6; DB 7; Length 600;
Best Local Similarity 82.6%; Pred. No. 88;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 AAAAATCCTGTAGGCTTTTCAGTG 25
Db 129 AATACTCCTGGGGGCTTCAGTG 107

RESULT 38
US-10-750-623-21331/c
; Sequence 21331, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 21331
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Bovine MMBT04380
US-10-750-623-21331

Query Match 66.4%; Score 16.6; DB 7; Length 600;
Best Local Similarity 82.6%; Pred. No. 88;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 AAAAATCCTGTAGGCTTTTCAGTG 25
Db 129 AATACTCCTGGGGGCTTCAGTG 107

RESULT 39
US-10-750-185-26142
; Sequence 26142, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922

; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 26142
; LENGTH: 781
; TYPE: DNA
; ORGANISM: Bovine 19866881217231
US-10-750-185-26142

Query Match 66.4%; Score 16.6; DB 7; Length 781;
Best Local Similarity 82.6%; Pred. No. 93;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AAAAATCCTGTAGGCTTTTCAGT 24
Db 621 AAAAATCCTGAAGGGTATCAGT 643

RESULT 40
US-10-750-623-26142
; Sequence 26142, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 26142
; LENGTH: 781
; TYPE: DNA
; ORGANISM: Bovine 19866881217231
US-10-750-623-26142

Query Match 66.4%; Score 16.6; DB 7; Length 781;
Best Local Similarity 82.6%; Pred. No. 93;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AAAAATCCTGTAGGCTTTTCAGT 24
Db 621 AAAAATCCTGAAGGGTATCAGT 643

RESULT 41
US-10-750-185-52730/c
; Sequence 52730, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 52730
; LENGTH: 914

Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAACTCTGTAGGCTTTTCAG 23
|||||
Db 303 AAAAAATCCTGTTGACTTGCAG 281

RESULT 46

US-10-750-623-44378/c
; Sequence 44378, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 44378
; LENGTH: 1496
; TYPE: DNA
; ORGANISM: Bovine 19866880525142
US-10-750-623-44378

Query Match 66.4%; Score 16.6; DB 7; Length 1496;
Best Local Similarity 82.6%; Pred. No. 1.1e+02;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAACTCTGTAGGCTTTTCAG 23
|||||
Db 303 AAAAAATCCTGTTGACTTGCAG 281

RESULT 47

US-10-750-185-52605/c
; Sequence 52605, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 52605
; LENGTH: 1507
; TYPE: DNA
; ORGANISM: Bovine 19866880525142
US-10-750-185-52605

Query Match 66.4%; Score 16.6; DB 7; Length 1507;
Best Local Similarity 82.6%; Pred. No. 1.1e+02;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AAAAACTCTGTAGGCTTTTCAGT 24

Db 303 AAAAAATCCTTTTGGCATTTCAGT 281
|||||

RESULT 48

US-10-750-623-52605/c
; Sequence 52605, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 52605
; LENGTH: 1507
; TYPE: DNA
; ORGANISM: Bovine 19866880525142
US-10-750-623-52605

Query Match 66.4%; Score 16.6; DB 7; Length 1507;
Best Local Similarity 82.6%; Pred. No. 1.1e+02;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AAAAACTCTGTAGGCTTTTCAGT 24
|||||
Db 303 AAAAAATCCTTTTGGCATTTCAGT 281

RESULT 49

US-10-750-185-50030/c
; Sequence 50030, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 50030
; LENGTH: 1695
; TYPE: DNA
; ORGANISM: Bovine 19866880411509
US-10-750-185-50030

Query Match 66.4%; Score 16.6; DB 7; Length 1695;
Best Local Similarity 82.6%; Pred. No. 1.1e+02;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AAAAACTCTGTAGGCTTTTCAGT 24
|||||
Db 215 AAAAACTCTGAACGCTTTTCAGT 193

```
RESULT 50
US-10-750-623-50030/c
; Sequence 50030, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KEER, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 50030
; LENGTH: 1695
; TYPE: DNA
; ORGANISM: Bovine 1986680411509
US-10-750-623-50030

Query Match      66.4%; Score 16.6; DB 7; Length 1695;
Best Local Similarity 82.6%; Pred. No. 1.1e+02;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      2  AAAAAGCTCCTGTAGGCTTTCAGT 24
Db      215 AAAAAGCTTCTGAACGCTTTCAC 193

Search completed: February 3, 2006, 16:20:07
Job time : 346.111 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 3, 2006, 10:53:23 ; Search time 863 Seconds
(without alignments)
1646.682 Million cell updates/sec

Title: US-10-719-900-25

Perfect score: 25

Sequence: 1 aaaaactgcattcccaatgttatg 25

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Databases :

GenEmbl:*

1: gb_ba:*

2: gb_in:*

3: gb_env:*

4: gb_om:*

5: gb_ov:*

6: gb_pat:*

7: gb_ph:*

8: gb_pr:*

9: gb_ro:*

10: gb_sts:*

11: gb_sy:*

12: gb_un:*

13: gb_vi:*

14: gb_htg:*

15: gb_pl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	23.4	93.6	170951	9 AC123552	AC123552 Mus muscu
2	23.4	93.6	228147	9 AC090652	AC090652 Mus muscu
3	22.4	89.6	3146	9 RRPCP3145	XG2840 R.rattus MR
4	22.4	89.6	3977	6 AX839774	AX839774 Sequence
5	22.4	89.6	3977	6 AX839887	AX839887 Sequence
6	22.4	89.6	3977	9 RATKV4	M6880 Rat neurona
7	22.4	89.6	217732	14 AC103001	AC103001 Rattus no
8	22.4	89.6	348540	14 AC132503	AC132503 Rattus no
9	21.8	87.2	270	6 AR270447	AR270447 Sequence
10	21.8	87.2	460	6 BD118543	BD118543 EST and e
11	21.8	87.2	460	6 AR422990	AR422990 Sequence
12	21.8	87.2	460	6 AX983684	AX983684 Sequence
13	21.8	87.2	1859	8 AX123157	AX123157 Homo sapi
14	21.8	87.2	115958	8 AC004736	AC004736 Human Chr
15	21.8	87.2	115971	8 AC124056	AC124056 Homo sapi
16	21.8	87.2	177142	8 AC004582	AC004582 Human Chr
17	21.8	87.2	177543	8 AC124078	AC124078 Homo sapi
18	20.8	83.2	109919	14 AP007327	AP007327 Lotus cor

19	20.4	81.6	1917	6 AR484315	AR484315 Sequence
20	20.4	81.6	1917	6 AX142745	AX142745 Sequence
21	20.4	81.6	3276	1 AF270200	AF270200 Staphyloc
22	20.4	81.6	3276	6 AR486154	AR486154 Sequence
23	20.4	81.6	3276	6 AX145518	AX145518 Sequence
C 24	20.4	81.6	110000	1 CP000029_20	Continuation (21 o
C 25	20.4	81.6	306698	1 AE016750	AE016750 Staphyloc
26	20.2	80.8	1382	5 BC099258	BC099258 Xenopus l
27	20.2	80.8	1731	5 BC077190	BC077190 Xenopus l
28	19.8	79.2	66122	8 AL732290	AL732290 Human DNA
29	19.8	79.2	148468	14 AL355882	AL355882 Homo sapi
C 30	19.8	79.2	148497	14 AC079592	AC079592 Homo sapi
31	19.8	79.2	185810	14 AC136052	AC136052 Rattus no
C 32	19.8	79.2	189389	8 CNS05TGF	AL355096 Human chr
33	19.8	79.2	211635	14 AC112763	AC112763 Rattus no
C 34	19.4	77.6	2124	5 BC045302	BC045302 Danio rer
C 35	19.4	77.6	46296	5 BX276107	BC276107 Zebrafish
36	19.4	77.6	58289	5 BX470142	BX470142 Zebrafish
37	19.4	77.6	140962	14 CR387991	CR387991 Danio rer
38	19.4	77.6	172243	9 AL606513	AL606513 Mouse DNA
C 39	19.4	77.6	185950	14 AL606486	AL606486 Homo sapi
C 40	19.4	77.6	228881	14 CR936412	CR936412 Danio rer
41	19.2	76.8	479	2 AY598742	AY598742 Trilobodr
C 42	19.2	76.8	694	10 BV660808	BV660808 S216P6047
C 43	19.2	76.8	959	10 BV524469	BV524469 G591P6048
44	19.2	76.8	988	10 BV543020	BV543020 sp143b08
C 45	19.2	76.8	3079	2 AB122062	AB122062 Crassosotr
C 46	19.2	76.8	39499	2 U41015	U41015 Caenorhabdi
C 47	19.2	76.8	110000	15 AP008209_256	Continuation (257
C 48	19.2	76.8	155293	14 AC154085	AC154085 Rhinolph
49	19.2	76.8	157381	8 AL360267	AL360267 Human DNA
50	19.2	76.8	163328	14 AC016190	AC016190 Homo sapi
51	19.2	76.8	164394	9 AC103388	AC103388 Mus muscu
C 52	19.2	76.8	164690	14 AC016846	AC016846 Homo sapi
C 53	19.2	76.8	166437	14 AC135133	AC135133 Rattus no
54	19.2	76.8	167427	15 AC139174	AC139174 Oryza sat
55	19.2	76.8	170756	5 BX248508	BX248508 Zebrafish
C 56	19.2	76.8	171245	14 CR762414	CR762414 Danio rer
57	19.2	76.8	173432	14 AC141166	AC141166 Rattus no
C 58	19.2	76.8	179144	5 BX511025	BX511025 Zebrafish
59	19.2	76.8	179651	8 AC005553	AC005553 Homo sapi
60	19.2	76.8	187298	8 AC018691	AC018691 Homo sapi
61	19.2	76.8	188004	14 AC113894	AC113894 Rattus no
62	19.2	76.8	189055	8 AC015724	AC015724 Homo sapi
63	19.2	76.8	198903	8 BS000070	BS000070 Pan trogl
64	19.2	76.8	200521	9 AL833804	AL833804 Mouse DNA
65	19.2	76.8	203959	8 AP001137	AP001137 Homo sapi
C 66	19.2	76.8	206136	8 AC068037	AC068037 Homo sapi
67	19.2	76.8	219117	9 AC129332	AC129332 Mus muscu
68	19.2	76.8	222594	14 AC124931	AC124931 Rattus no
69	19.2	76.8	224716	14 AC133852	AC133852 Rattus no
C 70	19.2	76.8	225959	14 CR925879	CR925879 Danio rer
C 71	19.2	76.8	232537	14 AC157781	AC157781 Mus muscu
C 72	19.2	76.8	238807	14 AC116254	AC116254 Rattus no
73	19.2	76.8	239423	14 AC096022	AC096022 Rattus no
C 74	19.2	76.8	248921	14 AC095538	AC095538 Rattus no
75	19.2	76.8	268898	14 AC109095	AC109095 Rattus no
76	19.2	76.8	340000	8 AP001681	AP001681 Homo sapi
77	19	76.0	159468	5 AL935040	AL935040 Zebrafish
78	19	76.0	170054	14 BX537266	BX537266 Danio rer
79	19	76.0	173631	5 BX537275	BX537275 Zebrafish
80	18.8	75.2	307	10 CFA411044	CFA411044 Canis fam
C 81	18.8	75.2	589	10 BV429969	BV429969 S237P6264
C 82	18.8	75.2	717	10 BV517218	BV517218 qdv77c02
C 83	18.8	75.2	729	5 CR406939	CR406939 Gallus ga
C 84	18.8	75.2	771	10 BV535492	BV535492 G591P6286
85	18.8	75.2	775	10 BV527407	BV527407 G591P6108
86	18.8	75.2	786	10 BV535377	BV535377 G591P6301
87	18.8	75.2	815	10 BV517667	BV517667 r1n39c05
C 88	18.8	75.2	856	10 BV512020	BV512020 qbn32d10
89	18.8	75.2	53542	6 AX247730	AX247730 Sequence
90	18.8	75.2	96387	14 HSU600024	AL080247 Homo sapi
C 91	18.8	75.2	109374	8 AP000952	AP000952 Homo sapi

VNKTEIENVRNGTQVRYREAEATEAFLTYIEGVCVVWFTEBFLMRVVFPCNKKVFIKN
SLNIIDFVAIILPFYLEVGLSGSSKAADVLGFLKVRVFRILIRIFKLTIRHFVGLRVL
GHTLRASNEFLLLIIIFALGVLIIFATMIYAEIRGAQPNPDSASEHFKNIPIGFW
WAVVTMTTLGYGDMYPOTWSGMLVGCALCAGVLTIAMPVPIVNNFGMYSLAMAKQ
KLPKKKKHIPRPOLGSPNYCKSVNSPHHSTQSDTCPLAQEETLEINRADSKLNGE
VAKAALANEDCPHIDQALTPDEGLPTFRSGTRERYGCPCLLSTGEYACPPGGMRKDL
CKESVIAKYMPTEAVRVVT"
3958. .3963

polyA_signal
ORIGIN

Query Match 89.6%; Score 22.4; DB 6; Length 3977;
Best Local Similarity 95.8%; Pred. No. 12;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AAAAACTGCATCCCAATGTTATG 25
|||||
Db 3915 AAAAACTGCATCCCAATGTTATG 3938

RESULT 5
AX839887
LOCUS AX839887 3977 bp DNA linear PAT 16-DEC-2003
DEFINITION Sequence 1 from Patent EP1348964.
ACCESSION AX839887
VERSION AX839887.1 GI:39978418
KEYWORDS
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus

REFERENCE
AUTHORS Cochran,S.Y., Yamagami,K.M. and Ohashi,Y.M.
TITLE Schizophrenia related gene
JOURNAL Patent: EP 1348964-A 1 01-OCT-2003;
Mitsubishi Pharma Corporation (JP)
FEATURES
source
1. .3977
/organism="Rattus norvegicus"
/mol_type="unassigned DNA"
/db_xref="taxon:10116"
1162. .2919
/note="unamed protein product"
/codon_start=1
/protein_id="CAF85433.1"
/db_xref="GI:39978419"
1162. .2919
/CDS

polyA_signal
ORIGIN

Query Match 89.6%; Score 22.4; DB 6; Length 3977;
Best Local Similarity 95.8%; Pred. No. 12;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AAAAACTGCATCCCAATGTTATG 25
|||||
Db 3915 AAAAACTGCATCCCAATGTTATG 3938

RESULT 6
RATKV4
LOCUS RATKV4 3977 bp mRNA linear ROD 27-APR-1993
DEFINITION Rat neuronal delayed rectifier K+ channel (K-V-4) mRNA, complete

cds.
M68880 M37845
M68880.1 GI:205106
neuronal delayed rectifier K+ channel.
KEYWORDS Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus
ORGANISM Rattus norvegicus

REFERENCE
AUTHORS Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidae; Murinae; Rattus.
1 (bases 1 to 3977)
Luneau,C.J., Williams,J.B., Marshall,J., Levitan,E.S., Oliva,C.,
Smith,J.S., Antanavage,J., Foidander,K., Stein,R.B., Swanson,R.,
Kaczmarek,L.K. and Buhrow,S.A.
TITLE Alternative splicing contributes to K+ channel diversity in the
mammalian central nervous system
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 88 (9), 3932-3936 (1991)
PUBMED 2023941
COMMENT Original source text: Rattus norvegicus (strain Sprague-Dawley)
brain cDNA to mRNA.

FEATURES
Location/Qualifiers
1. .3977
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/tissue_type="brain"
1. .3977
/gene="K-V-4"
/gene="K-V-4"
1. .1161
/gene="K-V-4"
1162. .2919
/gene="K-V-4"
/codon_start=1
/product="neuronal delayed rectifier K+ channel"
/protein_id="AAA41501.1"
/db_xref="GI:205107"
/translation="MGQDSESRIVNCGTRHQTYSRLTLPGLRLAWLAEPDAHS
HFDYPRADRFPPDRHPGVFAHILNRYRTGKLCPADVCGPLYEELAFWGDIDTDE
PCWMTYRQHRDAEALDSFGGAPLONSADADAGPGSDGDELEMTKRLALSDS
PDGREGFWRQPRIWALFEDPYSSRYARYVAFSLFFILVSIITFCLETHERNPI
VNKTEIENVRNGTQVRYREAEATEAFLTYIEGVCVVWFTEBFLMRVVFPCNKKVFIKN
SLNIIDFVAIILPFYLEVGLSGSSKAADVLGFLKVRVFRILIRIFKLTIRHFVGLRVL
WAVVTMTTLGYGDMYPOTWSGMLVGCALCAGVLTIAMPVPIVNNFGMYSLAMAKQ
KLPKKKKHIPRPOLGSPNYCKSVNSPHHSTQSDTCPLAQEETLEINRADSKLNGE
VAKAALANEDCPHIDQALTPDEGLPTFRSGTRERYGCPCLLSTGEYACPPGGMRKDL
CKESVIAKYMPTEAVRVVT"
2920. .3977
/gene="K-V-4"
3958. .3963
/gene="K-V-4"
polyA_signal
ORIGIN

Query Match 89.6%; Score 22.4; DB 9; Length 3977;
Best Local Similarity 95.8%; Pred. No. 12;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AAAAACTGCATCCCAATGTTATG 25
|||||
Db 3915 AAAAACTGCATCCCAATGTTATG 3938

RESULT 7
AC103001/c
LOCUS AC103001 217732 bp DNA linear HTG 13-MAY-2003
DEFINITION Rattus norvegicus clone CH230-185F19, WORKING DRAFT SEQUENCE, 3
unordered pieces.
ACCESSION AC103001
VERSION AC103001.8 GI:30581312
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GIRU
Center clone name: CH230-185F19
----- Summary Statistics
Assembly program: Atlas 3.0;
Consensus quality: 184743 bases at least Q40
Consensus quality: 185828 bases at least Q30
Consensus quality: 186720 bases at least Q20
Estimated insert size: 189327; sum-of-contigs estimation
Quality coverage: 9x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
* as soon as it is available and the accession number will
* be preserved.

1 215333: contig of 215333 bp in length
215334 215433: gap of unknown length
215334 216477: contig of 1044 bp in length
216478 216577: gap of unknown length
216578 217732: contig of 1155 bp in length.

----- Location/Qualifiers
1. 217732
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-185F19"
1. 1505
/note="wgs end extension
clone_end:T7"
2095. 4056
/note="wgs end extension
clone_end:T7"
complement(5859. 6364)
/note="clone boundary
clone_end:T7
site:ECORI
end sequence:BH263208"
complement(213207. 213763)
/note="clone boundary
clone_end:Sp5
site:ECORI
end sequence:BH263257"
215334. 215433
/estimated_length=unknown
216478. 216577
/estimated_length=unknown

Query Match 89.6%; Score 22.4; DB 14; Length 217732;
Best Local Similarity 95.8%; Pred. No. 9;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Cy 2 AAAAAGTCATCCCAATGTATG 25
|||||
Db 93910 AAAAAGTCATCCCAATGTATG 93887

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidae; Muridae; Murinae; Rattus.	
1 (bases 1 to 217732)	
Muzny, D. Marie, Metzker, M. Lee, Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Altschrocks, S., Amin, A., Anguiano, D., Anyalebechi, V., Ayagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biewalo, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Georegeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Haves, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpachy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensuewa, L., Louissegh, H., Lozada, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhney, S., McLeod, M. P., McNeill, T. Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwaokemele, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Poindestex, A., Popovic, D., Primus, E., Pu, L., L., Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, R., Rose, R., Ruiz, S. J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smajls, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villalana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczky, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstock, G. and Gibbs, R. A.	
Unpublished	
2 (bases 1 to 217732)	
Worley, K. C.	
Direct Submission	
Submitted (24-NOV-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA	
3 (bases 1 to 217732)	
Rat Genome Sequencing Consortium.	
Direct Submission	
Submitted (13-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA	
On May 13, 2003 this sequence version replaced gi:24942563.	
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated	

REFERENCE
AUTHORS

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT

```

RESULT 8
AC132503
LOCUS
DEFINITION
Rattus norvegicus clone CH230-425P24, *** SEQUENCING IN PROGRESS
***, 9 unordered pieces.
AC132503
AC132503.3 GI:25139153
HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
KEYWORDS
Rattus norvegicus (Norway rat)
SOURCE
Rattus norvegicus
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Murioidea; Muridae; Murinae; Rattus;
1 (bases 1 to 348540)
Muzny, D. Marie., Metzker, M. Lee., Abramson, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,
Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G.,
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
Gregg, G., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
Gunnarone, P., Haaland, W., Hamil, C., Henderson, N., Hamilton, K.,
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogue, M.,
Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jollivet, A.,
Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
Kowis, C., Kraat, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
Lorensuewa, L., Loulseged, H., Lozado, R. J., Lu, X., Ma, J.,
Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A.,
Mangum, B., Mapa, P., Martin, K., Martin, R., Martinez, E.,
Mawhney, S., McLeod, M. P., McNeill, T. Z., Meenen, E.,
Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
Morgan, M., Morris, K., Morris, S., Munitasa, M., Murphy, M., Nair, L.,
Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
Nwaokelemeh, O., Okwuonu, G., Olarnpusagoon, A., Pal, S., Parks, K.,
Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,
Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.,
Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R.,
Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J.,
Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,
Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smajs, D.,
Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,
Steinle, M., Strong, R., Sutton, A., Svatek, A., Taber, F., Taylor, C.,
Taylor, R., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,
Valas, R., Vera, V., Villalana, D., Waldron, L., Walker, B., Wang, J.,
Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
Williams, G., Willson, R., Wleczek, R., Wooden, H., Worley, K.,
Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O.,
Weinstock, G. and Gibbs, R. A.
Direct Submission
Unpublished
2 (bases 1 to 348540)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (01-SEP-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 348540)

```

AUTHORS TITLE JOURNAL

COMMENT

Rat Genome Sequencing Consortium.
 Direct Submission
 Submitted (20-NOV-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Nov 20, 2002 this sequence version replaced gi:23908470.
 The sequence in this assembly is a combination of BAC based reads
 and whole genome shotgun sequencing reads assembled using Atlas
 (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
 in the feature table below represents a scaffold in the Atlas
 assembly (a 'contig-scaffold'). Within each contig-scaffold,
 individual sequence contigs are ordered and oriented, and separated
 by sized gaps filled with Ns to the estimated size. The sequence
 may extend beyond the ends of the clone and there may be sequence
 contigs within a contig-scaffold that consist entirely of whole
 genome shotgun sequence reads. Both end sequences and whole genome
 shotgun sequence only contigs will be indicated in the feature
 table.

----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: http://www.hgsc.bcm.tmc.edu/
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: KBRQ
 Center clone name: CH230-425P24
 ----- Summary Statistics
 Assembly program: Phrap; version 0.990329
 Consensus quality: 295693 bases at least Q40
 Consensus quality: 301215 bases at least Q30
 Consensus quality: 304883 bases at least Q20
 Estimated insert size: 295566; sum-of-contigs estimation
 Quality coverage: 5x in Q20 bases; sum-of-contigs estimation

 * NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)
 * NOTE: This sequence may represent more than one clone.
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 9 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 31755: contig of 31755 bp in length
 * 31756 31855: gap of unknown length
 * 31856 60925: contig of 29070 bp in length
 * 60926 61025: gap of unknown length
 * 61026 64397: contig of 3372 bp in length
 * 64398 64497: gap of unknown length
 * 64498 130320: contig of 65823 bp in length
 * 130321 130420: gap of unknown length
 * 130421 146576: contig of 16156 bp in length
 * 146577 146676: gap of unknown length
 * 146677 157802: contig of 11126 bp in length
 * 157803 157902: gap of unknown length
 * 157903 346082: contig of 188180 bp in length
 * 346083 346182: gap of unknown length
 * 346183 347276: contig of 1094 bp in length
 * 347277 347376: gap of unknown length
 * 347377 348540: contig of 1164 bp in length.

FEATURES
 source
 1. 348540
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-425P24"
 11273..11359
 /note="clone_boundary
 clone_end:Sp6
 site:
 end_sequence:BZ211011"

misc_feature

TITLE JOURNAL REFERENCE AUTHORS TITLE JOURNAL

REFERENCE

```
misc_feature 23469..24939
/note="wgs contig"
misc_feature 24990..26503
/note="wgs contig"
gap 31756..31855
/estimated_length=unknown
misc_feature 31856..32991
/note="wgs contig"
misc_feature 33343..34933
/note="wgs contig"
gap 60926..61025
/estimated_length=unknown
misc_feature 61026..62237
/note="wgs contig"
gap 64398..64497
/estimated_length=unknown
misc_feature 102826..106931
/note="wgs contig"
gap 130321..130420
/estimated_length=unknown
misc_feature 130421..131444
/note="wgs contig"
gap 146577..146676
/estimated_length=unknown
misc_feature 156232..157802
/note="wgs contig"
gap 157803..157902
/estimated_length=unknown
misc_feature 157903..158937
/note="wgs contig"
misc_feature 161582..163160
/note="wgs contig"
misc_feature 184173..184811
/note="clone_boundary
clone_end:T7
site:
end_sequence:BZ211009"
misc_feature 343981..346082
/note="wgs_end_extension
clone_end:T7"
gap 346083..346182
/estimated_length=unknown
gap 347277..347376
/estimated_length=unknown

ORIGIN
Query Match 89.6%; Score 22.4; DB 14; Length 348540;
Best Local Similarity 95.8%; Pred. No. 8.7;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AAAAACTGCATCCCAATGTTATG 25
|||||
Db 237083 AAAAACTGCATCCCAATGTTATG 237106
|||||

RESULT 9
AR270447 AR270447 270 bp DNA linear PAT 10-APR-2003
LOCUS Sequence 1010 from patent US 6500938.
DEFINITION AR270447
ACCESSION AR270447
VERSION AR270447.1 GI:29701681
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 270)
AUTHORS Au-Young, J. and Sellhammer, J.J.
TITLE Composition for the detection of signaling pathway gene expression
JOURNAL Patent: US 6500938-A 1010 31-DEC-2002;
Incyte Genomics, Inc.; Palo Alto, CA;
WOX;
FEATURES
source Location/Qualifiers
1..270

/organism="unknown"
/mol_type="genomic DNA"

Query Match 87.2%; Score 21.8; DB 6; Length 270;
Best Local Similarity 92.0%; Pred. No. 28;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAACTGCATCCCAATGTTATG 25
|||||
Db 164 AAAAACTGCATCCCAATGTTATG 188
|||||

RESULT 10
BD118543 BD118543 460 bp DNA linear PAT 18-SEP-2002
LOCUS EST and encoded human protein.
DEFINITION BD118543
ACCESSION BD118543
VERSION BD118543.1 GI:23213453
KEYWORDS JP 2002010789-A/10620.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
REFERENCE 1 (bases 1 to 460)
AUTHORS Edwards, J.B.D.M., Jobert, S. and Giordano, J.E.
TITLE EST and encoded human protein
JOURNAL Patent: JP 2002010789-A 10620 15-JAN-2002;
GENSET CORP
COMMENT OS Homo sapiens (human)
PN JP 2002010789-A/10620
PD 15-JAN-2002
PF 07-AUG-2000 JP 2000280989
PR 05-AUG-1999 US 60/147499
PI JEAN BAPTISTE DUMAS MILNE EDWARDS, SEVELIN JOBERT, JEAN EVE PI
GIORDANO
PC C12N15/09, C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, PC
C12N1/21,
PC C12N5/10, C12P21/02, C12P21/08, C12Q1/68, C12N15/00, C12N5/00, PC
C12N15/00
CC EST and encoded human protein
FH Key Location/Qualifiers
FT source 1..460
FT /organism="Homo sapiens (human)".
FT Location/Qualifiers
1..460
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 87.2%; Score 21.8; DB 6; Length 460;
Best Local Similarity 92.0%; Pred. No. 27;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAACTGCATCCCAATGTTATG 25
|||||
Db 136 AAAAACTGCATCCCAATGTTATG 160
|||||

RESULT 11
AR422990 AR422990 460 bp DNA linear PAT 18-DEC-2003
LOCUS Sequence 14487 from patent US 6639063.
DEFINITION AR422990
ACCESSION AR422990
VERSION AR422990.1 GI:40178100
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 460)
AUTHORS Edwards, J.-B.D.M., Jobert, S. and Giordano, J.-Y.
```

```

TITLE      EST's and encoded human proteins
JOURNAL    Patent: US 6639063-A 14487 28-OCT-2003;
           Genset S.A.;
FEATURES   Location/Qualifiers
           1. .460
           /organism="unknown"
           /mol_type="genomic DNA"
ORIGIN
Query Match      87.2%; Score 21.8; DB 6; Length 460;
Best Local Similarity 92.0%; Pred. No. 27;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAACTGCATCCCAATGTTATG 25
    ||||| ||||| ||||| ||||| |||||
Db 136 AAAAACCTGCATGCCCAATGTTATG 160

RESULT 12
AX983684      460 bp      DNA      PAT 15-JAN-2004
LOCUS         Sequence 14487 from Patent EP1104808.
ACCESSION     AX983684
VERSION       AX983684.1 GI:40989824
KEYWORDS      Homo sapiens (human)
SOURCE        Homo sapiens
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
              Homnidae; Homo.
REFERENCE     1
AUTHORS       Dumas Milne Edwards,J.B., Jobert,S. and Giordano,J.Y.
TITLE         ESTs and encoded human proteins
JOURNAL      Patent: EP 1104808-A 14487 06-JUN-2001;
           Genset (FR)
FEATURES      Location/Qualifiers
           1. .460
           /organism="Homo sapiens"
           /mol_type="unassigned DNA"
           /db_xref="taxon:9606"
ORIGIN
Query Match      87.2%; Score 21.8; DB 6; Length 460;
Best Local Similarity 92.0%; Pred. No. 27;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAACTGCATCCCAATGTTATG 25
    ||||| ||||| ||||| ||||| |||||
Db 136 AAAAACCTGCATGCCCAATGTTATG 160

RESULT 13
AK123157      1859 bp      mRNA      linear      PRI 30-JAN-2004
LOCUS         Homo sapiens cDNA FLJ41162 fis, clone BRACE2039600, highly similar
           to VOLTAGE-GATED POTASSIUM CHANNEL PROTEIN KV3.1.
DEFINITION     AK123157
ACCESSION     AK123157.1 GI:34528632
VERSION       AK123157.1
KEYWORDS      oligo capping; fis (full insert sequence).
SOURCE        Homo sapiens
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
              Homnidae; Homo.
REFERENCE     1
AUTHORS       Ota,T., Suzuki,Y., Nishikawa,T., Otsuki,T., Sugiyama,T., Irie,R.,
           Wakamatsu,A., Hayashi,K., Sato,H., Nagai,K., Kimura,K., Makita,H.,
           Sekine,M., Oobayashi,M., Nishi,T., Shibahara,T., Tanaka,T.,
           Ishii,S., Yamamoto,M., Saito,K., Kawai,Y., Isono,Y., Nakamura,Y.,
           Nagahari,K., Murakami,K., Yasuda,T., Iwayanagi,T., Wagatsuma,M.,
           Shiratori,A., Sudo,H., Hosoiri,T., Kaku,Y., Kodaira,H., Kondo,H.,
           Sugawara,M., Takahashi,M., Kanda,K., Kanda,K., Yokoi,T., Furuya,T.,

```

```

Kikkawa,E., Omura,Y., Abe,K., Kamihara,K., Katsuta,N., Sato,K.,
Tanikawa,M., Yamazaki,M., Ninomiya,K., Ishibashi,T., Yamashita,H.,
Murakawa,K., Fujimori,K., Tanai,H., Kimata,M., Watanabe,M.,
Hiraoka,S., Chiba,Y., Ishida,S., Ono,Y., Takiguchi,S., Watanabe,S.,
Yasuda,M., Hotuta,T., Kusano,J., Kanehori,K., Takahashi-Fujii,A.,
Hara,H., Tanase,T., Nomura,Y., Togiya,S., Komai,F., Hara,R.,
Takeuchi,K., Arita,M., Imose,N., Musashino,K., Yuuki,H., Oshima,A.,
Sasaki,N., Aotsuka,S., Yoshikawa,Y., Matsunawa,H., Ichihara,T.,
Shichata,N., Sano,S., Moriya,S., Momiyama,H., Satoh,N., Takami,S.,
Terashima,Y., Suzuki,O., Nakagawa,S., Senoh,A., Mizoguchi,H.,
Goto,Y., Shimizu,F., Wakebe,H., Hishigaki,H., Watanabe,T.,
Sugiyama,A., Takemoto,M., Kawakami,B., Yamazaki,M., Watanabe,K.,
Kumagai,A., Itakura,S., Fukuzumi,Y., Fujimori,Y., Komiyama,M.,
Tashiro,H., Tanigami,A., Fujiwara,T., Ono,T., Yamada,K., Fujii,Y.,
Ozaki,K., Hirao,M., Ohmori,Y., Kawabata,A., Hikiji,T., Kobatake,N.,
Inagaki,H., Ikema,Y., Okamoto,S., Okitani,R., Kawakami,T.,
Noguchi,S., Itoh,T., Shigetake,K., Senba,T., Matsumura,K.,
Nakajima,Y., Mizuno,T., Morinaga,M., Sasaki,M., Togaishi,T.,
Oyama,M., Hata,H., Watanabe,M., Komatsu,T., Mizushima-Sugano,J.,
Satoh,T., Shirai,Y., Takahashi,Y., Nakagawa,K., Okumura,K.,
Nagase,T., Nomura,N., Kikuchi,H., Masuho,Y., Yamashita,R.,
Nakai,K., Yada,T., Nakamura,Y., Ohara,O., Isogai,T. and Sugano,S.
Complete sequencing and characterization of 21,243 full-length
human cDNAs
Nat. Genet. 36 (1), 40-45 (2004)
14702039
2
Kawakami,B., Sugiyama,A., Takemoto,M., Sugiyama,T., Irie,R.,
Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y.,
Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H.,
Watsu,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K.,
Wagatsuma,M., Murakawa,K., Kanehori,K., Takahashi-Fujii,A.,
Oshima,A., Suzuki,Y., Sugano,S., Nagahari,K., Masuho,Y., Nagai,K.
and Isogai,T.
NEDO human cDNA sequencing project
Unpublished
3 (bases 1 to 1859)
Isogai,T. and Yamamoto,J.
Direct Submission
Submitted (15-JUL-2003) Takao Isogai, FLJ Project (HRI Team); 2-6-7
Kazusa-Kamatari, Kisarazu, Chiba 292-0818, Japan
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction: Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: HRI and
RAB; annotation: HRI and RAB.
FEATURES      Location/Qualifiers
           1. .1859
           /organism="Homo sapiens"
           /mol_type="mRNA"
           /db_xref="taxon:9606"
           /clone="BRACE2039600"
           /tissue_type="cerebellum"
           /clone_lib="BRACE2"
           /note="cloning vector: pME18SFL3"
ORIGIN
Query Match      87.2%; Score 21.8; DB 8; Length 1859;
Best Local Similarity 92.0%; Pred. No. 24;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAACTGCATCCCAATGTTATG 25
    ||||| ||||| ||||| ||||| |||||
Db 1472 AAAAACCTGCATGCCCAATGTTATG 1496

RESULT 14
AC004736/c
LOCUS         AC004736      115958 bp      DNA      linear      PRI 03-JUL-1998

```


DEFINITION	Human Chromosome lip14.3 PAC clone pDJ1082L12 containing KNCN1 and MyoD, complete sequence.	FEATURES	Location/Qualifiers
ACCESSION	AC004736	source	1. .115958
VERSION	AC004736.1 GI:3282183		/organism="Homo sapiens"
KEYWORDS	HTG.		/mol_type="genomic DNA"
SOURCE	Homo sapiens (human)		/db_xref="taxon:9606"
ORGANISM	Homo sapiens	repeat_region	4409..4517
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.	repeat_region	/rpt_family="MIR"
REFERENCE	1 (bases 1 to 115958)	repeat_region	complement(19171..19291)
AUTHORS	Evans,G.A., Athanasiou,M., Aguayo,P., Armstrong,D., Basit,M., Buettner,J., Bumeister,R., Card,P., deSailboat,F., Dunn,J., English,C., Ethridge,S., Garner,H.R., Gee,V., Gordon,M., Gotway,G., Grant,O., Hahner,L., Harris,J., Lewis,E., Loo,H., Loo,K.N., McFarland,J., Miller,R., Newton,J., Osborne-Lawrence,S., Schageman,J., Schultz,R.A., Stimson,S., Syed,M. and Ward,T.	repeat_region	/rpt_family="Alu"
	HTGS Submission	repeat_region	complement(21561..21985)
TITLE	Unpublished	repeat_region	23678..23767
JOURNAL	2 (bases 1 to 115958)	repeat_region	/rpt_family="MIR"
AUTHORS	Evans,G.A., Athanasiou,M., Aguayo,P., Armstrong,D., Basit,M., Buettner,J., Bumeister,R., Card,P., deSailboat,C., Dunn,J., English,C., Ethridge,S., Garner,H.R., Gee,V., Gordon,M., Gotway,G., Grant,O., Hahner,L., Harris,J., Lewis,E., Loo,H., Loo,K.N., McFarland,J., Miller,R., Newton,J., Osborne-Lawrence,S., Schageman,J., Schultz,R.A., Stimson,S., Syed,M. and Ward,T.	repeat_region	28877..31250
	Direct Submission	repeat_region	/rpt_family="Tigger1"
TITLE	Submitted (23-MAY-1998) Genome Science & Technology Center, University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd, Dallas, TX 75235-8591, USA	repeat_region	complement(31777..31876)
JOURNAL	3 (bases 1 to 115958)	repeat_region	/rpt_family="MIR"
AUTHORS	Evans,G.A., Athanasiou,M., Aguayo,P., Armstrong,D., Basit,M., Buettner,J., Bumeister,R., Card,P., deSailboat,F., Dunn,J., English,C., Ethridge,S., Garner,H.R., Gee,V., Gordon,M., Gotway,G., Grant,O., Hahner,L., Harris,J., Lewis,E., Loo,H., Loo,K.N., Major,T., McFarland,J., Newton,J., Osborne-Lawrence,S., Schageman,J., Schultz,R.A., Stimson,S., Syed,M. and Ward,T.	repeat_region	complement(36023..36231)
	Direct Submission	repeat_region	/rpt_family="MIR"
TITLE	Submitted (01-JUL-1998) Genome Science & Technology Center, University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd, Dallas, TX 75235-8591, USA	repeat_region	complement(39712..39766)
JOURNAL	4 (bases 1 to 115958)	repeat_region	/rpt_family="MIR"
AUTHORS	Evans,G.A., Athanasiou,M., Aguayo,P., Armstrong,D., Basit,M., Buettner,J., Bumeister,R., Card,P., deSailboat,F., Dunn,J., English,C., Ethridge,S., Garner,H.R., Gee,V., Gordon,M., Gotway,G., Grant,O., Hahner,L., Harris,J., Lewis,E., Loo,H., Loo,K.N., Major,T., McFarland,J., Newton,J., Osborne-Lawrence,S., Schageman,J., Schultz,R.A., Stimson,S., Syed,M. and Ward,T.	repeat_region	complement(41971..42073)
	Direct Submission	repeat_region	/rpt_family="MIR"
TITLE	Submitted (03-JUL-1998) Genome Science & Technology Center, University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd, Dallas, TX 75235-8591, USA	repeat_region	45642..45752
JOURNAL	5 (bases 1 to 115958)	repeat_region	/rpt_family="MIR"
AUTHORS	Evans,G.A., Athanasiou,M., Aguayo,P., Armstrong,D., Basit,M., Buettner,J., Bumeister,R., Card,P., deSailboat,F., Dunn,J., English,C., Ethridge,S., Garner,H.R., Gee,V., Gordon,M., Gotway,G., Grant,O., Hahner,L., Harris,J., Lewis,E., Loo,H., Loo,K.N., Major,T., McFarland,J., Newton,J., Osborne-Lawrence,S., Schageman,J., Schultz,R.A., Stimson,S., Syed,M. and Ward,T.	repeat_region	57204..57337
	Direct Submission	repeat_region	/rpt_family="MIR"
TITLE	Submitted (01-JUL-1998) Genome Science & Technology Center, University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd, Dallas, TX 75235-8591, USA	repeat_region	59316..59729
JOURNAL	6 (bases 1 to 115958)	repeat_region	/rpt_family="L1"
AUTHORS	Evans,G.A., Athanasiou,M., Aguayo,P., Armstrong,D., Basit,M., Buettner,J., Bumeister,R., Card,P., deSailboat,F., Dunn,J., English,C., Ethridge,S., Garner,H.R., Gee,V., Gordon,M., Gotway,G., Grant,O., Hahner,L., Harris,J., Lewis,E., Loo,H., Loo,K.N., Major,T., McFarland,J., Newton,J., Osborne-Lawrence,S., Schageman,J., Schultz,R.A., Stimson,S., Syed,M. and Ward,T.	repeat_region	complement(60013..60386)
	Direct Submission	repeat_region	/rpt_family="THE1"
TITLE	Submitted (01-JUL-1998) Genome Science & Technology Center, University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd, Dallas, TX 75235-8591, USA	repeat_region	60398..60596
JOURNAL	7 (bases 1 to 115958)	repeat_region	/rpt_family="L1"
AUTHORS	Evans,G.A., Athanasiou,M., Aguayo,P., Armstrong,D., Basit,M., Buettner,J., Bumeister,R., Card,P., deSailboat,F., Dunn,J., English,C., Ethridge,S., Garner,H.R., Gee,V., Gordon,M., Gotway,G., Grant,O., Hahner,L., Harris,J., Lewis,E., Loo,H., Loo,K.N., Major,T., McFarland,J., Newton,J., Osborne-Lawrence,S., Schageman,J., Schultz,R.A., Stimson,S., Syed,M. and Ward,T.	repeat_region	60940..61228
	Direct Submission	repeat_region	/rpt_family="Alu"
TITLE	Submitted (01-JUL-1998) Genome Science & Technology Center, University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd, Dallas, TX 75235-8591, USA	repeat_region	complement(63635..63900)
JOURNAL	8 (bases 1 to 115958)	repeat_region	/rpt_family="Alu"
AUTHORS	Evans,G.A., Athanasiou,M., Aguayo,P., Armstrong,D., Basit,M., Buettner,J., Bumeister,R., Card,P., deSailboat,F., Dunn,J., English,C., Ethridge,S., Garner,H.R., Gee,V., Gordon,M., Gotway,G., Grant,O., Hahner,L., Harris,J., Lewis,E., Loo,H., Loo,K.N., Major,T., McFarland,J., Newton,J., Osborne-Lawrence,S., Schageman,J., Schultz,R.A., Stimson,S., Syed,M. and Ward,T.	repeat_region	complement(75510..75774)
	Direct Submission	repeat_region	/rpt_family="Alu"
TITLE	Submitted (03-JUL-1998) Genome Science & Technology Center, University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd, Dallas, TX 75235-8591, USA	repeat_region	complement(82106..82378)
JOURNAL	9 (bases 1 to 115958)	repeat_region	/rpt_family="Alu"
AUTHORS	Evans,G.A., Athanasiou,M., Aguayo,P., Armstrong,D., Basit,M., Buettner,J., Bumeister,R., Card,P., deSailboat,F., Dunn,J., English,C., Ethridge,S., Garner,H.R., Gee,V., Gordon,M., Gotway,G., Grant,O., Hahner,L., Harris,J., Lewis,E., Loo,H., Loo,K.N., Major,T., McFarland,J., Newton,J., Osborne-Lawrence,S., Schageman,J., Schultz,R.A., Stimson,S., Syed,M. and Ward,T.	repeat_region	complement(84150..84211)
	Direct Submission	repeat_region	/rpt_family="MLT1"
TITLE	Submitted (03-JUL-1998) Genome Science & Technology Center, University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd, Dallas, TX 75235-8591, USA	repeat_region	84727..85172
JOURNAL	10 (bases 1 to 115958)	repeat_region	/rpt_family="MLT1"
AUTHORS	Evans,G.A., Athanasiou,M., Aguayo,P., Armstrong,D., Basit,M., Buettner,J., Bumeister,R., Card,P., deSailboat,F., Dunn,J., English,C., Ethridge,S., Garner,H.R., Gee,V., Gordon,M., Gotway,G., Grant,O., Hahner,L., Harris,J., Lewis,E., Loo,H., Loo,K.N., Major,T., McFarland,J., Newton,J., Osborne-Lawrence,S., Schageman,J., Schultz,R.A., Stimson,S., Syed,M. and Ward,T.	repeat_region	85852..86134
	Direct Submission	repeat_region	/rpt_family="Alu"
TITLE	Submitted (03-JUL-1998) Genome Science & Technology Center, University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd, Dallas, TX 75235-8591, USA	repeat_region	86387..86565
JOURNAL	11 (bases 1 to 115958)	repeat_region	/rpt_family="Alu"
AUTHORS	Evans,G.A., Athanasiou,M., Aguayo,P., Armstrong,D., Basit,M., Buettner,J., Bumeister,R., Card,P., deSailboat,F., Dunn,J., English,C., Ethridge,S., Garner,H.R., Gee,V., Gordon,M., Gotway,G., Grant,O., Hahner,L., Harris,J., Lewis,E., Loo,H., Loo,K.N., Major,T., McFarland,J., Newton,J., Osborne-Lawrence,S., Schageman,J., Schultz,R.A., Stimson,S., Syed,M. and Ward,T.	repeat_region	complement(88437..88514)
	Direct Submission	repeat_region	/rpt_family="MIR"
TITLE	Submitted (03-JUL-1998) Genome Science & Technology Center, University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd, Dallas, TX 75235-8591, USA	repeat_region	complement(90686..90752)
JOURNAL	12 (bases 1 to 115958)	repeat_region	/rpt_family="MIR"
AUTHORS	Evans,G.A., Athanasiou,M., Aguayo,P., Armstrong,D., Basit,M., Buettner,J., Bumeister,R., Card,P., deSailboat,F., Dunn,J., English,C., Ethridge,S., Garner,H.R., Gee,V., Gordon,M., Gotway,G., Grant,O., Hahner,L., Harris,J., Lewis,E., Loo,H., Loo,K.N., Major,T., McFarland,J., Newton,J., Osborne-Lawrence,S., Schageman,J., Schultz,R.A., Stimson,S., Syed,M. and Ward,T.	repeat_region	91042..91163
	Direct Submission	repeat_region	/rpt_family="MIR"
TITLE	Submitted (03-JUL-1998) Genome Science & Technology Center, University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd, Dallas, TX 75235-8591, USA	repeat_region	94274..94419
JOURNAL	13 (bases 1 to 115958)	repeat_region	/rpt_family="MIR"
AUTHORS	Evans,G.A., Athanasiou,M., Aguayo,P., Armstrong,D., Basit,M., Buettner,J., Bumeister,R., Card,P., deSailboat,F., Dunn,J., English,C., Ethridge,S., Garner,H.R., Gee,V., Gordon,M., Gotway,G., Grant,O., Hahner,L., Harris,J., Lewis,E., Loo,H., Loo,K.N., Major,T., McFarland,J., Newton,J., Osborne-Lawrence,S., Schageman,J., Schultz,R.A., Stimson,S., Syed,M. and Ward,T.	repeat_region	complement(95470..95617)
	Direct Submission	repeat_region	/rpt_family="MLT1"
TITLE	Submitted (03-JUL-1998) Genome Science & Technology Center, University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd, Dallas, TX 75235-8591, USA	repeat_region	complement(97484..97917)
JOURNAL	14 (bases 1 to 115958)	repeat_region	/rpt_family="L1"
AUTHORS	Evans,G.A., Athanasiou,M., Aguayo,P., Armstrong,D., Basit,M., Buettner,J., Bumeister,R., Card,P., deSailboat,F., Dunn,J., English,C., Ethridge,S., Garner,H.R., Gee,V., Gordon,M., Gotway,G., Grant,O., Hahner,L., Harris,J., Lewis,E., Loo,H., Loo,K.N., Major,T., McFarland,J., Newton,J., Osborne-Lawrence,S., Schageman,J., Schultz,R.A., Stimson,S., Syed,M. and Ward,T.	repeat_region	105586..105657
	Direct Submission	repeat_region	/rpt_family="MIR"
TITLE	Submitted (03-JUL-1998) Genome Science & Technology Center, University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd, Dallas, TX 75235-8591, USA	repeat_region	105673..105884
JOURNAL	15 (bases 1 to 115958)	repeat_region	/rpt_family="MIR"
AUTHORS	Evans,G.A., Athanasiou,M., Aguayo,P., Armstrong,D., Basit,M., Buettner,J., Bumeister,R., Card,P., deSailboat,F., Dunn,J., English,C., Ethridge,S., Garner,H.R., Gee,V., Gordon,M., Gotway,G., Grant,O., Hahner,L., Harris,J., Lewis,E., Loo,H., Loo,K.N., Major,T., McFarland,J., Newton,J., Osborne-Lawrence,S., Schageman,J., Schultz,R.A., Stimson,S., Syed,M. and Ward,T.	repeat_region	complement(106369..106641)
	Direct Submission	repeat_region	/rpt_family="Alu"
TITLE	Submitted (03-JUL-1998) Genome Science & Technology Center, University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd, Dallas, TX 75235-8591, USA	repeat_region	complement(107302..107596)
JOURNAL	16 (bases 1 to 115958)	repeat_region	/rpt_family="Alu"
AUTHORS	Evans,G.A., Athanasiou,M., Aguayo,P., Armstrong,D., Basit,M., Buettner,J., Bumeister,R., Card,P., deSailboat,F., Dunn,J., English,C., Ethridge,S., Garner,H.R., Gee,V., Gordon,M., Gotway,G., Grant,O., Hahner,L., Harris,J., Lewis,E., Loo,H., Loo,K.N., Major,T., McFarland,J., Newton,J., Osborne-Lawrence,S., Schageman,J., Schultz,R.A., Stimson,S., Syed,M. and Ward,T.	repeat_region	109188..109458
	Direct Submission	repeat_region	/rpt_family="Alu"

MAPPED CLONE OVERLAP: pDJ59m18 and 6-109h6.
 IMPORTANT: This submission contains the entire insert of clone pDJ1082L12. pDJ1082L12 comes from a PAC library constructed at the Roswell Park Cancer Institute by the Pieter de Jong group. This clone has been finished according to strict quality criteria and attempts have been made to resolve all base calling problems such as compressions and repetitive elements. The expected Phred/Phrap calculated errors/10kb is 0.39. In addition, this sequence has been finished such that 99.9% of consensus base calls consist of either double-stranded coverage or 2 types of labeling chemistry on one strand.

[illegible]

```
repeat_region      /rpt_family="L1ME3A"
4970_..5179
/rpt_family="L1MD3"
5180_..5490
/rpt_family="AluSx"
5491_..5609
/rpt_family="L1MD3"
5609_..5845
/rpt_family="L1ME3A"
complement(6442..6757)
/rpt_family="AluSx"
8356_..8671
/rpt_family="AluSx"
8717_..8725
/note="<30 qual SNGL region"
8735_..8770
/note="<30 qual SNGL region"
8786_..8790
/note="<30 qual SNGL region"
9317_..9609
/rpt_family="AluSc"
complement(10070..10405)
/rpt_family="MIR"
complement(11554..11859)
/rpt_family="L1R16C"
12223_..12374
/rpt_family="MIR"
13250_..13353
/rpt_family="MIR3"
complement(14122..14358)
/rpt_family="MIR"
14653_..14758
/rpt_family="MIR"
16324_..17416
/rpt_family="L2"
complement(17963..18020)
/rpt_family="MLTIE2"
18027_..18705
/rpt_family="L1MB5"
complement(18706..19168)
/rpt_family="MLTIE2"
19383_..19447
/rpt_family="L2"
complement(20015..20094)
/rpt_family="L2"
20108_..20563
/rpt_family="MLTIF"
21080_..21223
/rpt_family="L3"
complement(21504..21725)
/rpt_family="MIR"
complement(22444..22513)
/rpt_family="MIR3"
22970_..23148
/rpt_family="G-rich"
24021_..24120
/rpt_family="L2"
complement(24730..24934)
/rpt_family="MIR"
24947_..25068
/rpt_family="L2"
25137_..25277
/rpt_family="MIR"
25441_..25539
/rpt_family="MIR"
26018_..26087
/rpt_family="L2"
27434_..27568
/rpt_family="MIR"
complement(28003..28114)
/rpt_family="MIR"
28115_..28608
/rpt_family="MLT1H"

repeat_region      complement(28622..29055)
/rpt_family="L1MC1"

Query Match      87.2%; Score 21.8; DB 8; Length 115971;
Best Local Similarity 92.0%; Pred. No. 18;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAACTGCATCCCAATGTTATG 25
|||||
Db 109770 AAAAAGCTGCATGCCCAATGTTATG 109794

RESULT 16
AC004582 177142 bp DNA linear PRI 23-JUL-1998
LOCUS Human Chromosome 11p14.3 PAC clone pDJ59m18, complete sequence.
DEFINITION AC004582
ACCESSION AC004582.1 GI:3337307
VERSION HTG.
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM
REFERENCE
AUTHORS
1 (bases 1 to 177142)
Evans,G.A., Athanasiou,M., Aguayo,P., Armstrong,D., Basit,M.,
Buettnner,J., Bumeister,R., Card,P., desailboat,C., Dunn,J.,
English,C., Ethridge,S., Garner,H.R., Gee,V., Gordon,M., Gotway,G.,
Grant,O., Hahner,L., Joslin,J., Lewis,E., Loo,H., Loo,K.N.,
Mcfarland,J., Miller,R., Newton,J., Osborne-Lawrence,S.,
Schageman,J., Schultz,R.A., Stimson,S., Syed,M. and Ward,T.
HTGS Submission
Unpublished
TITLE
JOURNAL
REFERENCE
AUTHORS
2 (bases 1 to 177142)
Evans,G.A., Athanasiou,M., Aguayo,P., Armstrong,D., Basit,M.,
Buettnner,J., Bumeister,R., Card,P., desailboat,C., Dunn,J.,
English,C., Ethridge,S., Garner,H.R., Gee,V., Gordon,M., Gotway,G.,
Grant,O., Hahner,L., Joslin,J., Lewis,E., Loo,H., Loo,K.N.,
Mcfarland,J., Miller,R., Newton,J., Osborne-Lawrence,S.,
Schageman,J., Schultz,R.A., Stimson,S., Syed,M. and Ward,T.
Direct Submission
Submitted (17-APR-1998) Genome Science & Technology Center,
University of Texas Southwestern Medical Center, 5323 Harry Hines
Blvd, Dallas, TX 75235-8591, USA
TITLE
JOURNAL
REFERENCE
AUTHORS
3 (bases 1 to 177142)
Evans,G.A., Athanasiou,M., Aguayo,P., Armstrong,D., Basit,M.,
Buettnner,J., Bumeister,R., Card,P., desailboat,C., Dunn,J.,
English,C., Ethridge,S., Garner,H.R., Gee,V., Gordon,M., Gotway,G.,
Grant,O., Hahner,L., Joslin,J., Lewis,E., Loo,H., Loo,K.N.,
Major,T., McFarland,J., Newton,J., Osborne-Lawrence,S.,
Schageman,J., Schultz,R.A., Stimson,S., Syed,M. and Ward,T.
Direct Submission
Submitted (23-JUL-1998) Genome Science & Technology Center,
University of Texas Southwestern Medical Center, 5323 Harry Hines
Blvd, Dallas, TX 75235-8591, USA
COMMENT
On Jul 23, 1998 this sequence version replaced gi:3057010.
IMPORTANT: This submission contains the entire insert of clone
pDJ59m18. pDJ59m18 comes from a PAC library constructed at the
Roswell Park Cancer Institute by the Pieter de Jong group. This
clone has been finished according to strict quality criteria and
attempts have been made to resolve all base calling problems such
as compressions and repetitive elements. The expected Phred/Phrap
calculated errors/10kb is 0.40. In addition, this sequence has
been finished such that 99.9% of consensus base calls consist of
either double-stranded coverage or 2 types of labeling chemistry on
one strand.
NOTE: Extent of pDJ59m18 overlap with mapped clones; SP6 PAC
Cloning end: 1. 8775. pDJ1082Li2.
Further information regarding the map of this region or annotation
of pDJ59m18 can be found at http://gestec.swmed.edu/chromos2.htm.
CHROMOSOMAL LOCUS: This PAC clone comes from the Chromosome 11p14.3
Usher(USH1C) syndrome region mapped between STSs D11S1310 and
1115A14.
```

MARKER CONFIRMATION: STS D11S1310 and SHGC-19182 sequence confirmed
MAPPED CLONE OVERLAP: pDJ1082L12 and 6-130a9.

```

QY      1  AAAAACTGCATCCCAATGTTATG 25
Db      2589 AAAAACTGCATCCCAATGTTATG 2613

RESULT 17
LOCUS   AC124078
DEFINITION Homo sapiens chromosome 11, clone RP1-59M18, complete sequence.
ACCESSION AC124078
VERSION   AC124078.5
KEYWORDS  HTG.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
REFERENCE 1 (bases 1 to 177543)
AUTHORS  Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLES   Homo sapiens chromosome 11, clone RP1-59M18
JOURNAL  Unpublished
REFERENCE 2 (bases 1 to 177543)
AUTHORS  Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L.,
Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,
Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., Fitzgerald,M., FitzHugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I.,
Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K.,
Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Lindblad-Toh,K.,
Liu,G., Maclean,C., Macdonald,P., Major,J., Marquis,N.,
Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Meldrim,J.,
Meneus,L., Mihova,T., Mienga,V., Murphy,T., Naylor,J., Nguyen,C.,
Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C.,
Rogov,P., Roman,J., Rosetti,M., Roy,A., Santosa,R., Schauer,S.,
Schupback,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Testaye,S.,
Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (09-JUN-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE 3 (bases 1 to 177543)
AUTHORS  Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B.,
Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J.,
Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
Liu,G., Maclean,C., Macdonald,P., Major,J., Matthews,C.,
McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mienga,V.,
Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Roy,A., Schauer,S., Schupback,R., Seaman,S., Severy,P.,
Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
Testaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (30-AUG-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE 4 (bases 1 to 177543)
AUTHORS  Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B.,
Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J.,
Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
Liu,G., Maclean,C., Macdonald,P., Major,J., Matthews,C.,
McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mienga,V.,
Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Roy,A., Schauer,S., Schupback,R., Seaman,S., Severy,P.,
Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
Testaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (06-DEC-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT  On Dec 6, 2002 this sequence version replaced gi:22549795.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence submissions@genome.wi.mit.edu
----- Project Information
Center project name: L27396
Center clone name: 59_M18
-----
FEATURES             Location/Qualifiers
     source            1..177543
                        /organism="Homo sapiens"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:9606"
                        /chromosome="11"
                        /map="11"
                        /clone="RP1-59M18"
                        /clone_lib="RPIC Human PAC library 1"
     repeat_region     1863..1886
                        /rpt_family="AT rich"
     repeat_region     complement(3634..3771)
                        /rpt_family="MIR3"
                        4136..4177
     unsure            /note="single clone coverage"
                        complement(4256..4380)
                        /rpt_family="MIR"
                        4499..4581
                        /note="single clone coverage"
                        complement(5386..5528)
     repeat_region     /rpt_family="L2"
                        6596..6692
                        /rpt_family="MIR3"
                        8588..9349
     repeat_region     /rpt_family="L2"
                        complement(9361..9434)
                        /rpt_family="MIR"
                        9746..9941
     repeat_region     /rpt_family="Alusg/x"
                        complement(10213..10423)
     repeat_region     /rpt_family="L1ME3"
                        complement(10510..10713)
     repeat_region     /rpt_family="L1ME3"
                        complement(10757..11016)
     repeat_region     /rpt_family="L1ME3"
                        complement(11029..11524)
     repeat_region     /rpt_family="L1ME"
                        11530..11782
     repeat_region     /rpt_family="MIR"
                        complement(12563..12712)
     repeat_region     /rpt_family="MIR"
                        complement(12733..12808)

```

/rpt_family="MRR58"
complement(12812..12908)
/rpt_family="MRR58"
14061..14296
/rpt_family="MLT1J"
complement(14325..14464)
/rpt_family="MIR"
complement(14494..14718)
/rpt_family="MIR"
complement(15635..15755)
/rpt_family="L2"
15772..15834
/rpt_family="LMC/D"
16257..16628
/rpt_family="MSTD"
complement(16876..16947)
/rpt_family="L2"
16990..17652
/rpt_family="L2"
complement(18156..18220)
/rpt_family="MIR"
complement(18680..18890)
/rpt_family="MIR"
complement(18959..19375)
/rpt_family="L3"
complement(20085..20389)
/rpt_family="LMC1"
complement(20423..20838)
20839..21147
/rpt_family="ALUY"
complement(21148..21607)
/rpt_family="L2"
22409..22949
/rpt_family="L2"
23384..23562
/rpt_family="MIR"
24404..24429
/rpt_family="MIR"
24495..24662
/rpt_family="MIR"
25101..25270
/rpt_family="MIR3"
complement(25457..25572)
/rpt_family="MIR"
25625..25802
/rpt_family="MIR"
complement(26265..26497)
/rpt_family="MIR"
complement(28127..28548)
/rpt_family="MSTB"
complement(29435..29481)
/rpt_family="MIR"
30280..30355
/rpt_family="L2"
complement(30376..30470)
/rpt_family="MIR3"
30555..30664
/rpt_family="L2"
complement(31042..31207)
/rpt_family="MIR3"

Query Match 87.2%; Score 21.8; DB 8; Length 177543;
Best Local Similarity 92.0%; Pred. No. 18;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AAAAAGCTGATCCCAATGTTATG 25
Db 2593 AAAAAGCTGATCCCAATGTTATG 2617

RESULT 18
AP007327

LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AP007327 109919 bp DNA linear HTG 28-DEC-2004
Lotus corniculatus var. japonicus chromosome 3 clone LjT24C17, ***
SEQUENCING IN PROGRESS ***, 17 unordered pieces.

AP007327 GI:56805643
HTG; HTGS PHASE1.
Lotus corniculatus var. japonicus (Lotus japonicus)
Lotus corniculatus var. japonicus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Lotaeae;
Lotus.

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

1
Kaneko, T., Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S.
Structural Analysis of a Lotus japonicus Genome. XI. Sequence
Features and Mapping of Nine hundred twenty-one TAC Clones
Unpublished
2 (bases 1 to 109919)
Sato, S.
Direct Submission
Submitted (26-OCT-2004) Shusei Sato, Kazusa DNA Research Institute,
Department of Plant Gene Research; 2-6-7 Kazusa-kamatari, Kisarazu,
Chiba, 292-0818, Japan (E-mail:ssato@kazusa.or.jp,
URL:http://www.kazusa.or.jp/, Tel:81-438-52-3935(ex.2337),
Fax:81-438-52-3934)

COMMENT

* NOTE: This is a 'working draft' sequence. It currently
* consists of 17 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 4679: contig of 4679 bp in length
* 4680 4779: gap of unknown length
* 4780 6066: contig of 1287 bp in length
* 6067 6166: gap of unknown length
* 6167 8356: contig of 2190 bp in length
* 8357 8456: gap of unknown length
* 8457 11229: contig of 2773 bp in length
* 11230 11329: gap of unknown length
* 11330 14861: contig of 3532 bp in length
* 14862 14961: gap of unknown length
* 14962 17757: contig of 2796 bp in length
* 17758 17857: gap of unknown length
* 17858 22893: contig of 4836 bp in length
* 22894 22793: gap of unknown length
* 22794 27233: contig of 4440 bp in length
* 27234 27333: gap of unknown length
* 27334 31851: contig of 4518 bp in length
* 31852 38853: contig of 6902 bp in length
* 38854 38953: gap of unknown length
* 38954 45565: contig of 6612 bp in length
* 45566 45665: gap of unknown length
* 45666 52253: contig of 6588 bp in length
* 52254 52353: gap of unknown length
* 52354 58364: contig of 6011 bp in length
* 58365 58464: gap of unknown length
* 58465 66744: contig of 8280 bp in length
* 66745 66844: gap of unknown length
* 66845 84478: contig of 17634 bp in length
* 84479 84578: gap of unknown length
* 84579 108308: contig of 23730 bp in length
* 108309 108408: gap of unknown length
* 108409 109919: contig of 1511 bp in length.

FEATURES
source

1. 109919
/organism="Lotus corniculatus var. japonicus"
/mol_type="genomic DNA"
/variety="japonicus"
/db_xref="taxon:34305"
/chromosome="3"

```

/clone="LjT24C17"
/clone_lib="LjT library"
/note="TAC clone:TW0406, synonym:Lotus japonicus"
4680..4779
gap /estimated_length=unknown
6067..6166
gap /estimated_length=unknown
8357..8456
gap /estimated_length=unknown
11230..11329
gap /estimated_length=unknown
14862..14961
gap /estimated_length=unknown
17758..17857
gap /estimated_length=unknown
22694..22793
gap /estimated_length=unknown
27234..27333
gap /estimated_length=unknown
31852..31951
gap /estimated_length=unknown
38854..38953
gap /estimated_length=unknown
45566..45665
gap /estimated_length=unknown
52254..52353
gap /estimated_length=unknown
58365..58464
gap /estimated_length=unknown
66745..66844
gap /estimated_length=unknown
84479..84578
gap /estimated_length=unknown
108309..108408
gap /estimated_length=unknown

ORIGIN
Query Match 83.2%; Score 20.8; DB 14; Length 109919;
Best Local Similarity 91.7%; Pred. No. 55;
Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AAAAAGTCATCCCAATGTTATG 25
|||||
Db 60479 AAAAAGTCATCCCAATGTTATG 60502

RESULT 19
AR484315
LOCUS AR484315 1917 bp DNA linear PAT 14-MAY-2004
DEFINITION Sequence 1467 from patent US 6703492.
ACCESSION AR484315
VERSION AR484315.1 GI:47247418
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS Kimmerly,W.J.
TITLE Staphylococcus epidermidis nucleic acids and proteins
JOURNAL Patent: US 6703492-A 1467 09-MAR-2004;
Smithline Beecham Corporation; Philadelphia, PA
FEATURES
source
1..1917
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 81.6%; Score 20.4; DB 6; Length 1917;
Best Local Similarity 95.5%; Pred. No. 1.1e+02;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 AAAAAGTCATCCCAATGTTAT 24
|||||

```

```

Db 1441 AAAAAGTCATCCCAATGTTAT 1462

RESULT 20
AX142745
LOCUS AX142745 1917 bp DNA linear PAT 31-MAY-2001
DEFINITION Sequence 1467 from Patent WO0134809.
ACCESSION AX142745
VERSION AX142745.1 GI:14282244
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS Kimmerly,W.J.
TITLE Staphylococcus epidermidis nucleic acids and proteins
JOURNAL Patent: WO 0134809-A 1467 17-MAY-2001;
GLAXO GROUP LIMITED (GB)
FEATURES
source
1..1917
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/notes="synthetic nucleic acid sequence"

ORIGIN
Query Match 81.6%; Score 20.4; DB 6; Length 1917;
Best Local Similarity 95.5%; Pred. No. 1.1e+02;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 AAAAAGTCATCCCAATGTTAT 24
|||||
Db 1441 AAAAAGTCATCCCAATGTTAT 1462

RESULT 21
AF270200
LOCUS AF270200 3276 bp DNA linear BCT 01-AUG-2000
DEFINITION Staphylococcus epidermidis strain SRI clone step.1053f03 genomic
sequence.
ACCESSION AF270200
VERSION AF270200.1 GI:9624109
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS Kimmerly,W.J., Taylor,J.David., Nelsen,A.J., Godlevski,M.M.,
Rubino,M.A., Nelson,F.J., Rivers,P.R., Torruella-Miller,I.,
Listenbee,S., Ashanti,C., Altshuller,G., Mammo,L., Shepherd,N.S.,
Fuchs,R., Fleming,T., Guan,X., Du,L., Cain,D.H., Miller,G.S. and
Furdon,P.J.
TITLE Transposon-mediated sequencing of the Staphylococcus epidermidis
genome
JOURNAL Unpublished
REFERENCE
AUTHORS Taylor,J.David., Kimmerly,W.J., Nelsen,A.J., Godlevski,M.M.,
Rubino,M.A., Nelson,F.J., Rivers,P.R., Torruella-Miller,I.,
Listenbee,S., Ashanti,C., Altshuller,G., Mammo,L., Shepherd,N.S.,
Fuchs,R., Fleming,T., Guan,X., Du,L., Cain,D.H., Miller,G.S. and
Furdon,P.J.
TITLE Direct Submission
JOURNAL Submitted (22-MAY-2000) Departments of Genomic Sciences and
Bioinformatics, Genetics Directorate, Glaxo Wellcome, Inc., 5 Moore
Drive, Research Triangle Park, North Carolina 27709-3398, USA
FEATURES
source
1..3276
/organism="Staphylococcus epidermidis"
/mol_type="genomic DNA"
/db_xref="taxon:1282"
/clone="step.1053f03"

ORIGIN

```

RESULT 24


```
/product="trNA-Val"
complement (451..522)
/product="trNA-Glu"
complement (524..598)
/product="trNA-Asn"
complement (612..726)
/product="5S ribosomal RNA"
complement (886..3713)
/product="23S ribosomal RNA"
complement (2718..2831)
/locus_tag="SE1755"
complement (2718..2831)
/locus_tag="SE1755"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAO05396.1"
/db_xref="GI:27316221"
/translations="MMIIGRALFGRGAPLGFNSDKLRMPINLTWSEHG"
complement (3977..5530)
/product="16S ribosomal RNA"
complement (5985..7052)
/locus_tag="SE1756"
complement (5985..7052)
/locus_tag="SE1756"
/codon_start=1
/transl_table=11
/product="ATP-binding Mrp-like protein"
/protein_id="AAO05397.1"
/db_xref="GI:27316222"
/translations="MLTVNQKEIVGALKDPIIDVLRSEGIQVDSIKDINHVSVK
VMAOLGGQQLDQWAIKALKENGANTVIRFELPSEVVERVYIGKSEKPTIEE
LLSQNPVERISASGKGVGKSTVAVNLVALARSKKVLGVADADYGFSPVDMGI
DERPGIDKEIIPVERHGVKVIISMAFFVENAPVIRGPMGLKMLNFTFTEQWGBLD
YLLDLDPGGTGDVLDVHSLPSSKEIIVTTPHTAAFAVARAGAMAKTEHTILGVI
ENMSYFESKTEKGVYFGKGGKGLSDELTQLFAELPLEQTPWNPDPSPSIYQSD
DRGLGYTLARVIASTOKQ"
complement (7203..8642)
/locus_tag="SE1757"
complement (7203..8642)
/locus_tag="SE1757"
/codon_start=1
/transl_table=11
/product="multidrug resistance protein"
/protein_id="AAO05398.1"
/db_xref="GI:27316223"
/translations="MASYELSKTKRNLIVSMILISAFVAILNQTLNTALPHIMRELN
INENTSOGLITGFWLVNGVMIPLTAYLMDKVKTRPLYLGAMGSLIGSIVAALPNFG
VLMARVIOAGVAGILMPLMQFTLFTLFSKEHGRFAMGLAGLVIQAPAIQPTFTGLI
IDNASWRPVFIVIGIALVTEIFGVVISISYNTTKETLDRSVIYSTLFGGLMLYAF
SSAGNLGFSNPVLCVSLISLIIIGIFVKRQIISNPLNLKIFNNKIFCFSTISMI
IMLSMGFALLIPLYQNALGLSALLSGVLMPGAILINGMSVFTGPKDYKQPRPLI
FTGFIILISCTFLCLKVDTSYTLIVIAIRMAFVSLMMPINTTGINALKTEDIS
HGTAIMNFGVMAGSLGTALMTFMSIGCAQWVSSSEHASKEMIQRQSVAVGVDSFA
LVTVFVILAFVLAIFKEPNHLTHNSRKV"
complement (8912..9382)
/locus_tag="SE1758"
complement (8912..9382)
/locus_tag="SE1758"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAO05399.1"
/db_xref="GI:27316224"
/translations="MMKFLNKSXHLVLTIVLTIFVISGAILFTLFGFLYGLSRIL
IYHLGDFSNKGYDNLIVYGVSIVLGYFTLFSIEHLMDFYKKNLKNPYQFQINFH
LISIVTIMPYFIVHLHYVNVNHFVIMIIIGFLFVCKEVFPESKNLNKK"
complement (9570..10907)
/locus_tag="SE1759"
complement (9570..10907)
/locus_tag="SE1759"
/codon_start=1
/transl_table=11
/product="multidrug transporter"
/protein_id="AAO05400.1"
/db_xref="GI:27316225"
/translations="MNLKSIITVMALILIMFMAALETISIISLALPTIKNSLNAGNLVS
LVTVVFALVIANPIVGLMSRPKIIIVAVVGVLLFALGSLMSGLSOTFTFLIISRT
VQFGAGYMSLSQIVPKLAEIPLRYKIMGVSGVWGSIIGPLGGAILEFASWH
WLFYINIPAIIVAILVLMTHFHPDETQVQSRPDKGLIIFYIFIALMGLGNQHH
IIFNIFSILALAVLWLLFKIENSIEQPLFPKTFNFIISLVIFITDILLAIITLMGNL
YIPVYQKLSLPSQSGFVFPPLSVAMITLNFNAKTEAHFTRKTLXICSPFVLVLS
SLIMFGLKLLPLLIAFVAVFAGLSFGYIYTKDSIVVQBEETSPKNMKMMSFPVALLKL
GSSVGSTIMGYMYALNVGLFGSNLHNVLGLVLIIVAVCLIVMMNTLYKSTIQS"
complement (11146..11829)
/locus_tag="SE1760"
complement (11146..11829)
/locus_tag="SE1760"
/codon_start=1
/transl_table=11
/product="hemolysin III"
/protein_id="AAO05401.1"
/db_xref="GI:27316226"
/translations="MSQSKRKNDSSVETFKDIIPLTFGEIIGNAASHGAAALLTLFI
LPYAAVHGFNGGTLSESISVSVVISIFMFISSITIHSMONNTSHKYLIRIIDHSMI
VVAISGTVPVLLSVVGGWGLVLTILLGTLTGILYKSIATKVNHRLSLIYLVWVG
WVGIIPLPIIIMRTSWWFPIFFLGLSYTICANFVAQKNKPYFHMWHIFIVLASFL
HMGIFPYFM"
complement (11848..13035)
/locus_tag="SE1761"
complement (11848..13035)
/locus_tag="SE1761"
/codon_start=1
/transl_table=11
/product="UDP-N-acetylglucosamine pyrophosphorylase"
/protein_id="AAO05402.1"
/db_xref="GI:27316227"
/translations="MLDKNLEKYNQEHLYVEYKLSMSEKNALDEKVDQLNLAIEIQD
LQDLVFNKRTIDDVSSSEVKEYKRSKLRNEERHTYEQKVEAIEENGEPVALLMAGS
QTRQLYKGGPKGSFEIEGSLFELQARQLIRKEETGHTINWTIMTSDINHKTIEYF
KQKHYFNDAHIIHFQKQDNIVALSEEGKVLNRDGHIMETPNGNGGVFKSLKAGVYL
DKMQDDHVKYIIFLNNIDNVKVLDPPLFAGFTVQSKDITSKTIQPKDSSESVRLNVV
DKDVTLEYSELDTDIVNQFNANIGIHAFKLGFIITSAVDRELPHYLAIKQLQLDEN
FGVERPTLKFELFYDIFRYGTSFVTLQVPREBEFSPLNKKEGKDSVHTATEDLKRK
DLI"
complement (13090..13599)
/locus_tag="SE1762"
complement (13090..13599)
/locus_tag="SE1762"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAO05403.1"
/db_xref="GI:27316228"
/translations="MTGKTHASCGFLVGAIITQYFHTDIFTSTISVILSVISSILPDI
CHTQSKIGRRPLTSFPIRLGHRFTFSLHFIIGISFLLYFIQTQPMYVMAIVIGM
FSHVLIIDITPKGVKLLYPLPFNINVSPIHFKTGGLVDVSLATALSVGAIYTLQPYLN
TMHWYWLK"
13776..14036
/locus_tag="SE1763"
13776..14036
/locus_tag="SE1763"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAO05404.1"
/db_xref="GI:27316229"
/translations="MAWTVKQNDNEVRIQWRVADIKIPNNEIKNVTTQDDIHAVPEEN
GKEISIGTSFGKTNVLIDTDQHLIITYQNDQKVNELTK"
14189..14503
/locus_tag="SE1764"
14189..14503
/locus_tag="SE1764"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
```


ACCESSION	BC077190	BC077190.1	GI:50416502	
VERSION				
KEYWORDS				
SOURCE	Xenopus laevis	(African clawed frog)		
ORGANISM	Xenopus laevis			
REFERENCE				
AUTHORS	Klein,S.L., Strauberg,R.L., Wagner,L., Pontius,J., Clifton,S.W. and Richardson,P.			
TITLE	Genetic and genomic tools for Xenopus research: The NIH Xenopus Initiative			
JOURNAL	Dev. Dyn.	225 (4),	384-391 (2002)	
PUBLISHED	12454917			
REFERENCE				
AUTHORS	2 (bases 1 to 1731)			
	Krausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shemen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettman,M., Madan,A., Rodrigues,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalius,D.E., Schnerch,A., Schein,J.E., Jones,S.D. and Marra,M.A.			
	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences			
TITLE	Proc. Natl. Acad. Sci. U.S.A.	99 (26),	16899-16903 (2002)	
JOURNAL	12477932			
PUBLISHED	3 (bases 1 to 1731)			
REFERENCE				
AUTHORS	Klein,S. and Gerhard,D.S.			
TITLE	Direct Submission			
JOURNAL	Submitted (19-JUL-2004) National Institutes of Health, Xenopus Gene Collection (XGC), National Institute of Child Health and Human Development, 6100 Executive Boulevard, Room 4B01, Rockville, MD 20892-7510, USA			
REMARK	NIH-MGC Project			
COMMENT	Contact: XGC help desk Email: cgapbs-t@mail.nih.gov Tissue Procurement: Dr. Igor David cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML) DNA Sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada info@bcgc.bc.ca			
	Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield, Andy Chan, Steve S. Chand, William Chow, Alison Cloutier, Ruth Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao, Kim MacDonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Morin, Teika Olson, Diana Palmquist, Anca Petrescu, Anna Liisa Prahbu, Parvaneh Saedi, JR Santos, Angelique Schnerch, Ursula Skalska, Duane Smalius, Jeff Scott, Miranda Tsai, George Yang, Jacquie Schein, Asim Siddiqui, Rob Holt, Marco Marra.			
	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LIML at: http://image.llnl.gov Series: IRAC Plate: 149 Row: 0 Column: 24 This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, Similarity but not identity to protein.			
FEATURES	1..1731			
source	/organism="Xenopus laevis" /mol_type="mRNA"			

/db_xref="taxon:8355"	
/clone="IMAGE:4057027"	
/tissue type="Lung, adult Xenopus"	
/clone_lib="NICHD_XGC_Lul"	
/lab_host="DH10B"	
/note="Vector: pCMV-SPORT6"	
<1..1731	
/gene="LOC445845"	
/db_xref="GeneID:445845"	
<1..746	
/gene="LOC445845"	
/codon_start=3	
/product="LOC445845 protein"	
/protein_id="AAH77190.1"	
/db_xref="GI:50416503"	
/db_xref="GeneID:445845"	
/translation="TATCTIVTVRGSRSLGTFCCLARKKPKGAMRFFITFLCQIAFI LUTVSEEQKTEBEQVKIEVHLHPSECSQKSGDGLVNAHYDAYLAQNTKLKSGR ATKDGHKFWVLGVGVQIKGLDIALMCMCSGEKKRAIIPPSLAYGAGHDKIPPNATL MFEILYGINRGPRAEAFYQIDLNNKQLSKDEISHYLTEFFKDGKQKQDPVSHGTI LTDI1PHKNDRDEDFISAKENYVYGHDEL"	
ORIGIN	
Query Match	80.8%; Score 20.2; DB 5; Length 1731;
Best Local Similarity	88.0%; Pred. No. 1.4e+02;
Matches	22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY	1 AAAAACTGCATCCCAATGTTATG 25
Db	1311 AAAAACTGCATACAGATGTTATG 1335
RESULT 28	
AL732290	
LOCUS	AL732290 Human DNA sequence from clone CTD-2338C19 on chromosome X, complete sequence.
DEFINITION	AL732290.4 GI:21531515
ACCESSION	AL732290
VERSION	HTG.
KEYWORDS	Homo sapiens (human)
SOURCE	Homo sapiens
ORGANISM	Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi; Eukaryota; Metazoa; Chordata; Euarthontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 66122)
AUTHORS	Whitehead,S.
TITLE	Direct Submission
JOURNAL	Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vega@sanger.ac.uk
COMMENT	Clone requests: clonerequest@sanger.ac.uk On Jun 21, 2002 this sequence version replaced gi:21217881. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome X, constructed by the Sanger Centre Chromosome X Mapping Group. Further information can be found at http://www.sanger.ac.uk/HGP/ChrX CTD-2338C19 is from the Caltech genomic sperm BAC library D. VECTOR: pBelOBAC11
	----- Genome Center Center: Wellcome Trust Sanger Institute Center code: SC Web site: http://www.sanger.ac.uk Contact: vega@sanger.ac.uk ----- This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such

as compressions and repeats; all regions were covered by at least one subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

FEATURES

source
Location/Qualifiers
1..66122
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="CTD-2338C19"
/clone_lib="CIT-HSP-D1"

ORIGIN

Query Match 79.2%; Score 19.8; DB 8; Length 66122;
Best Local Similarity 91.3%; Pred. No. 1.7e+02;
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 AAAACTGCATCCCAATGTTATG 25
||||| ||||||| |||||
Db 14935 AAAACTGAATCCCAATATTATG 14957

RESULT 29

AL355882 148468 bp DNA linear HTG 31-JAN-2002
LOCUS Homo sapiens chromosome X clone RP3-479K6, 14 unordered pieces.
ACCESSION AL355882
VERSION AL355882.4 GI:9909022
KEYWORDS HTG; HTGS PHASE1; HTGS_CANCELLED.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.

REFERENCE

1
McLay, K.
Direct Submission
Submitted (21-JAN-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Aug 24, 2000 this sequence version replaced gi:9797390.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: dj1185F7
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: plasmid; L08752; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 140943 bases at least Q40
Consensus quality: 143751 bases at least Q30
Consensus quality: 145662 bases at least Q20
Insert size: 147168; sum-of-contigs
Insert size: 132352; 17.0% error; agarose-fp
Quality coverage: 3.82x in Q20 bases; sum-of-contigs Quality
coverage: 4.24x in Q20 bases; agarose-fp

COMMENT

* NOTE: This is a 'working draft' sequence. It currently
* consists of 14 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 3189: contig of 3189 bp in length
* 3190 3289: gap of 100 bp
* 3290 9680: contig of 6391 bp in length
* 9681 9780: gap of 100 bp
* 9781 12692: contig of 2912 bp in length
* 12693 12792: gap of 100 bp

ORIGIN

Query Match 79.2%; Score 19.8; DB 14; Length 148468;
Best Local Similarity 91.3%; Pred. No. 1.6e+02;
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 3 AAAACTGCATCCCAATGTTATG 25
||||| ||||||| ||||||| |||||

* 12793 19581: contig of 6789 bp in length
* 19582 19681: gap of 100 bp
* 19682 36219: contig of 16538 bp in length
* 36220 36319: gap of 100 bp
* 36320 46785: contig of 10466 bp in length
* 46786 46885: gap of 100 bp
* 46886 62406: contig of 15521 bp in length
* 62407 62506: gap of 100 bp
* 62507 97792: contig of 35286 bp in length
* 97793 97892: gap of 100 bp
* 97893 109555: contig of 11663 bp in length
* 109556 109655: gap of 100 bp
* 109656 112376: contig of 2721 bp in length
* 112377 112476: gap of 100 bp
* 112477 119534: contig of 7058 bp in length
* 119535 119635: gap of 100 bp
* 119635 123375: contig of 3741 bp in length
* 123376 123476: gap of 100 bp
* 123476 141241: contig of 17766 bp in length
* 141242 141341: gap of 100 bp
* 141342 148468: contig of 7127 bp in length.

FEATURES

source

Location/Qualifiers
1..148468
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="X"
/clone="RP3-479K6"
/clone_lib="RPCL-3"

misc_feature

1..3189
/note="assembly fragment:00217"
fragment chain:1"
3290..9680
/note="assembly fragment:00951"
fragment chain:1"

misc_feature

clone_end:T7
vector_side:left"
9781..12692
/note="assembly fragment:00104"
fragment chain:2"

misc_feature

12793..19581
/note="assembly fragment:00296"
fragment chain:2"
19682..36219
/note="assembly fragment:00323"

misc_feature

36320..46785
/note="assembly fragment:00381"
46886..62406
/note="assembly fragment:00449"

misc_feature

62507..97792
/note="assembly fragment:00590"
97893..109555
/note="assembly fragment:00631"

misc_feature

109656..112376
/note="assembly fragment:00720"
112477..119534
/note="assembly fragment:00986"

misc_feature

119635..123375
/note="assembly fragment:01008"
123476..141241
/note="assembly fragment:01141"

misc_feature

141342..148468
/note="assembly fragment:01102"
clone_end:SP6
vector_side:right"

Db 81449 AAAACTGAATCCCAATATTATG 81471

RESULT 30
AC079592/c
LOCUS
DEFINITION AC079592 148497 bp DNA linear HTG 08-JAN-2003
Homo sapiens chromosome 3 clone RP11-75415, WORKING DRAFT SEQUENCE,
25 unordered pieces.

ACCESSION AC079592
VERSION AC079592.8 GI:20335660
KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.

REFERENCE 1 (bases 1 to 148497)
AUTHORS Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Albrooke,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,
Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,
Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
Bunay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hogue,M., Holloway,C., Hollins,B.,
Homesi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
Kratochvil,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H.,
Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
Mashayekh,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
Nguyen,N., Nickerson,E., Nwokwenkwo,S., Ogih,M., Okwuonu,G.,
Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
Peterson,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
Rivers,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savery,G.,
Scherer,S., Scott,G., Shen,H., Shoohtari,N., Sisson,I.,
Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,
Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
Umanik,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K.,
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.

Direct Submission
Unpublished
2 (bases 1 to 148497)
Worley,K.C.

Direct Submission
Unpublished
Submitted (04-SEP-2000) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 148497)
Worley,K.C.

Direct Submission
Unpublished
Submitted (08-JAN-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Apr 28, 2002 this sequence version replaced gi:18449584.

----- Genome Center

Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: HBBP
Center clone name: RP11-75415
----- Summary Statistics
Sequencing vector: M13;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 121512 bases at least Q40
Consensus quality: 129899 bases at least Q30
Consensus quality: 135781 bases at least Q20
Estimated insert size: 144188; sum-of-contigs estimation
Quality coverage: 3x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 25 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 2325: contig of 2325 bp in length
2326 2425: gap of unknown length
2426 4893: contig of 2468 bp in length
4894 4894: gap of unknown length
4894 8632: contig of 3639 bp in length
8633 8732: gap of unknown length
8733 11069: contig of 2337 bp in length
11070 11189: gap of unknown length
11189 13945: contig of 2776 bp in length
13945 14045: gap of unknown length
14045 17046: contig of 3001 bp in length
17046 17146: gap of unknown length
17146 19481: contig of 2335 bp in length
19481 19581: gap of unknown length
19581 22448: contig of 2867 bp in length
22448 22548: gap of unknown length
22548 25891: contig of 3343 bp in length
25891 25991: gap of unknown length
25991 30431: contig of 4440 bp in length
30431 30531: gap of unknown length
30531 34280: contig of 3749 bp in length
34280 34380: gap of unknown length
34380 37568: contig of 3188 bp in length
37568 43299: contig of 5631 bp in length
43299 43399: gap of unknown length
43399 50957: contig of 7558 bp in length
50957 51057: gap of unknown length
51057 57906: contig of 6849 bp in length
57906 58006: gap of unknown length
58006 63428: contig of 5422 bp in length
63428 63528: gap of unknown length
63528 69051: contig of 5523 bp in length
69051 76910: contig of 7759 bp in length
76910 77010: gap of unknown length
77010 83900: contig of 6890 bp in length
83900 84000: gap of unknown length
84000 91544: contig of 7544 bp in length
91544 91644: gap of unknown length
91644 101701: contig of 10057 bp in length
101701 101801: gap of unknown length
101801 110735: contig of 8934 bp in length
110735 110835: gap of unknown length
110835 123862: contig of 13027 bp in length
123862 123962: gap of unknown length

* 123963 135193: contig of 11231 bp in length
 * 135194 135293: gap of unknown length
 * 135294 148497: contig of 13204 bp in length.

FEATURES

Location/Qualifiers

1. .148497

/organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="3"
 /clone="RP11-75415"

2326. .2425

gap /estimated_length=unknown

4894. .4993

gap /estimated_length=unknown

8633. .8732

gap /estimated_length=unknown

11070. .11169

gap /estimated_length=unknown

13946. .14045

gap /estimated_length=unknown

17047. .17146

gap /estimated_length=unknown

19482. .19581

gap /estimated_length=unknown

22449. .22548

gap /estimated_length=unknown

25892. .25991

gap /estimated_length=unknown

30432. .30531

gap /estimated_length=unknown

34281. .34380

gap /estimated_length=unknown

37569. .37668

gap /estimated_length=unknown

43300. .43399

gap /estimated_length=unknown

50958. .51057

gap /estimated_length=unknown

57907. .58006

gap /estimated_length=unknown

63429. .63528

gap /estimated_length=unknown

69052. .69151

gap /estimated_length=unknown

76911. .77010

gap /estimated_length=unknown

83901. .84000

gap /estimated_length=unknown

91545. .91644

gap /estimated_length=unknown

101702. .101801

gap /estimated_length=unknown

110736. .110835

gap /estimated_length=unknown

123863. .123962

gap /estimated_length=unknown

135194. .135293

gap /estimated_length=unknown

ORIGIN

Query Match 79.2%; Score 19.8; DB 14; Length 148497;

Best Local Similarity 91.3%; Pred. No. 1.6e+02;

Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 AAAAAGTCATCCCAATGTTATG 25

||||| ||||| ||||| ||||| |||||

Db 51902 AAAAAGTCATCCCAATGTTATG 51880

RESULT 31

AC136052

LOCUS

DEFINITION

AC136052 Rattus norvegicus clone CH230-292B24, *** SEQUENCING IN PROGRESS

linear HTG 29-OCT-2002

***, 75 unordered pieces.

AC136052

VERSION AC136052.1 GI:24417923

KEYWORDS

HTG: HTGS_PHASE1. (Norway rat)

SOURCE

ORGANISM

Rattus norvegicus

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

Sciurognathi; Muridae; Murinae; Rattus.

1 (bases 1 to 185810)

Muzny,D.Marie., Metzker,M.Lee., Abramson,S., Adams,C., Alder,J.,

Allen,C., Allien,H., Alsbrooks,S., Amin,A., Anguiano,D.,

Anylebechechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,

Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,

Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,

Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,

Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,

Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,

Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,

Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,

Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,

Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,

Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,

Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,

Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,

Georgievski,B., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,

Guaratri,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,

Harvey,I., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,

Hernandez,R., Hines,S., Hladun,S.B., Hodgson,A., Hogue,M.,

Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,

Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,

Karpachy,S., Kelly,S., Khan,Z., King,L., King,L., Kovar,C.,

Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,

Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,

Lorensuwa,L., Louissegh,H., Lozano,R.J., Lu,X., Ma,J.,

Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,

Mangum,B., Mapua,P., Martin,K., Martin,R., Martinec,E.,

Mawhney,S., McLeod,M., McNeill,T., Meenen,E., Milosavljevic,A.,

Miner,G., Minja,E., Montemayor,J., Moore,S., Morgan,M., Morris,K.,

Morris,S., Munidasa,M., Murphy,M., Nair,L., Nankervis,C., Neal,D.,

Newton,N., Nguyen,N., Norris,S., Nwakoelmech,O., Okwuonu,G.,

Olarnunsgoon,A., Pal,S., Parks,K., Pasternak,S., Paul,H.,

Perez,A., Perez,L., Pfannkuch,C., Plopper,F., Poindexter,A.,

Popovic,D., Primus,E., Pu,L., Puazo,M., Quiroz,J., Rachlin,E.,

Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y.,

Reuter,M., Richards,S., Riggs,F., Rives,C., Rodkey,T., Rojas,A.,

Rose,M., Rose,R., Ruiz,S.J., Sanders,W., Savery,G., Scherer,S.,

Scott,G., Shattman,S., Shen,H., Shetty,J., Shvartsbeyn,A.,

Sisson,I., Sitter,C.D., Smajs,D., Sneed,A., Sodergren,E.,

Song,X.-Z., Sorelle,R., Sosa,J., Steimle,M., Strong,R., Sutton,A.,

Swatek,A., Tabor,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S.,

Tingey,A., Trejos,Z., Umani,K., Valas,R., Vera,V., Villaseana,D.,

Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J.,

Warren,R., Wei,X., White,F., Williams,G., Willson,R., Wlarczyk,R.,

Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S.,

Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,J., Zhou,X.,

Zhao,S., Dunn,D., von Niederhausern,A., Weiss,R., Smith,D.R.,

Holt,R.A., Smith,H.O., Weinstein,G. and Gibbs,R.A.

Direct Submission

Unpublished

2 (bases 1 to 185810)

Rat Genome Sequencing Consortium.

Submitted (29-OCT-2002) Human Genome Sequencing Center, Department

of Molecular and Human Genetics, Baylor College of Medicine, One

Baylor Plaza, Houston, TX 77030, USA

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: KDDT

Center clone name: CH230-292B24
 ----- Summary Statistics
 Sequencing vector: Plasmid;
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.990329
 Consensus quality: 142822 bases at least Q40
 Consensus quality: 147110 bases at least Q30
 Consensus quality: 150075 bases at least Q20

 * NOTE: Estimated insert size may differ from sequence length
 * (see http://www.bgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 75 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence.
 * as soon as it is available and the accession number will
 * be preserved.

1 1330: contig of 1330 bp in length
 * 1331 1430: gap of unknown length
 * 1431 2700: contig of 1270 bp in length
 * 2701 2800: gap of unknown length
 * 2801 3940: contig of 1140 bp in length
 * 3941 4040: gap of unknown length
 * 4041 5156: contig of 1116 bp in length
 * 5157 5256: gap of unknown length
 * 5257 6859: contig of 1603 bp in length
 * 6860 6959: gap of unknown length
 * 6960 8417: contig of 1457 bp in length
 * 8417 8516: gap of unknown length
 * 8517 9808: contig of 1292 bp in length
 * 9809 9909: gap of unknown length
 * 9909 11040: contig of 1131 bp in length
 * 11040 11139: gap of unknown length
 * 11140 12169: contig of 1030 bp in length
 * 12170 12269: gap of unknown length
 * 12270 14012: contig of 1743 bp in length
 * 14013 14112: gap of unknown length
 * 14113 15810: contig of 1698 bp in length
 * 15811 15910: gap of unknown length
 * 15911 17216: contig of 1306 bp in length
 * 17217 17316: gap of unknown length
 * 17317 19178: contig of 1862 bp in length
 * 19179 19278: gap of unknown length
 * 19279 20521: contig of 1243 bp in length
 * 20522 20621: gap of unknown length
 * 20622 21622: contig of 1001 bp in length
 * 21623 21722: gap of unknown length
 * 21723 22909: contig of 1187 bp in length
 * 22910 23009: gap of unknown length
 * 23010 25670: contig of 2661 bp in length
 * 25671 25770: gap of unknown length
 * 25771 27137: contig of 1367 bp in length
 * 27138 27237: gap of unknown length
 * 27238 29209: contig of 1972 bp in length
 * 29210 29309: gap of unknown length
 * 29310 31130: contig of 1821 bp in length
 * 31131 31230: gap of unknown length
 * 31231 32812: contig of 1582 bp in length
 * 32813 32912: gap of unknown length
 * 32913 34538: contig of 1626 bp in length
 * 34539 34638: gap of unknown length
 * 34639 35647: contig of 1009 bp in length
 * 35648 37822: contig of 2075 bp in length
 * 37823 37922: gap of unknown length
 * 37923 39275: contig of 1353 bp in length
 * 39276 39376: gap of unknown length
 * 39377 41609: contig of 2234 bp in length
 * 41610 42936: contig of 1227 bp in length
 * 42937 43036: gap of unknown length

* 43037 44087: contig of 1051 bp in length
 * 44088 44187: gap of unknown length
 * 44188 45431: contig of 1244 bp in length
 * 45432 45531: gap of unknown length
 * 45532 47879: contig of 2348 bp in length
 * 47880 47979: gap of unknown length
 * 47980 49432: contig of 1453 bp in length
 * 49433 49532: gap of unknown length
 * 49533 52083: contig of 2551 bp in length
 * 52084 52183: gap of unknown length
 * 52184 53926: contig of 1743 bp in length
 * 53927 54026: gap of unknown length
 * 54027 56161: contig of 2135 bp in length
 * 56162 56261: gap of unknown length
 * 56262 57990: contig of 1729 bp in length
 * 57991 58090: gap of unknown length
 * 58091 60844: contig of 2754 bp in length
 * 60845 60944: gap of unknown length
 * 60945 63300: contig of 2356 bp in length
 * 63301 63400: gap of unknown length
 * 63401 64854: contig of 1454 bp in length
 * 64855 64954: gap of unknown length
 * 64955 67199: contig of 2244 bp in length
 * 67200 67298: gap of unknown length
 * 67299 69764: contig of 2466 bp in length
 * 69765 69864: gap of unknown length
 * 69865 72035: contig of 2171 bp in length
 * 72036 72135: gap of unknown length
 * 72136 74158: contig of 2023 bp in length
 * 74159 74258: gap of unknown length
 * 74259 76503: contig of 2245 bp in length
 * 76504 76604: gap of unknown length
 * 76605 79168: contig of 2565 bp in length
 * 79169 79268: gap of unknown length
 * 79269 81982: contig of 2714 bp in length
 * 81983 82082: gap of unknown length
 * 82083 84203: contig of 2121 bp in length
 * 84204 84303: gap of unknown length
 * 84304 87139: contig of 2836 bp in length
 * 87140 87239: gap of unknown length
 * 87240 89111: contig of 2472 bp in length
 * 89112 89811: gap of unknown length
 * 89812 92229: contig of 2418 bp in length
 * 92230 92329: gap of unknown length
 * 92330 94448: contig of 2119 bp in length
 * 94449 94548: gap of unknown length
 * 94549 96815: contig of 2267 bp in length
 * 96816 96915: gap of unknown length
 * 96916 100471: contig of 3556 bp in length
 * 100472 100571: gap of unknown length
 * 100572 103523: contig of 2952 bp in length
 * 103524 103623: gap of unknown length

Query Match 79.2%; Score 19.8; DB 14; Length 185810;
 Best Local Similarity 91.3%; Pred. No. 1.6e+02;
 Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAACTGCATCCCACTGTTA 23
 DB 94723 AAAAACTGCATCCCACTTTA 94745

RESULT 32
 CNS05TCF/c
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

CNS05TCF 189389 bp DNA linear PRI 04-MAY-2001
 Human chromosome 14 DNA sequence BAC R-168L7 of library RCI-11
 from chromosome 14 of Homo sapiens (Human), complete sequence.
 AL355096
 AL355096.4 GI:13992190
 HTG.
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominae; Homo
1 (bases 1 to 189389)
Heilig.R., Petit.J.L., Vico,V., Dasilva,C., Robert,C., Wincker,P.,
Brotier,P., Cattolico.L., Barbe,V., Pelletier,E., Artiguenave,F.,
Levy,M., Eckenberg,R., Bruls,T., deBerardinis,V., Cruaud,C.,
Gyapay,G., Saurin,W. and Weissbach,J.
Sequencing of the human chromosome 14
Unpublished
2 (bases 1 to 189389)
Genoscope.
Direct Submission
Submitted (04-MAY-2001) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqrefgenoscope.cns.fr
- Web : www.genoscope.cns.fr)
On May 8, 2001 this sequence version replaced gi:11876112.
----- Genome Center
Center: Genoscope / Centre National de Sequencage
Center code: GS
Web site: http://www.genoscope.cns.fr/
Contact: SeqRef@genoscope.cns.fr

The following BAC sequence is oriented from the T7 to the SP6 end.
Upstream BAC (overlapping the T7 end) : R-1029J19 (AC=AL049836)
Downstream BAC (overlapping the SP6 end) : C-2561P5 -----
Summary Statistics
Assembly program: Phrap; version 2.0
Quality coverage: 8.44x in Q20 bases; sum-of-contigs

Overall quality chart :
Range : Bases
0 : 3
1 - 9 : 14
10 - 19 : 60
20 - 29 : 176
30 - 39 : 699
40 - 49 : 4072
50 - 59 : 8246
60 - 69 : 7529
70 - 79 : 13558
80 - 89 : 37982
90 - 99 : 117050

Percentage of bases with a quality value >= 40 : 99 %.

Location/Qualifiers
1. 189389
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="14"
/clone_lib="RPCT-11"
/clone="R-168L7"
67559..67694
/note="matching EMBL:AI005270
R4db:RH102424
dbSTS:STS69961
Identified using the e-PCR software (G. Schuler)"

FEATURES
source
STS
ORIGIN
Query Match 79.2%; Score 19.8; DB 8; Length 189389;
Best Local Similarity 91.3%; Pred. No. 1.6e+02;
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 AAAAACTGCATCCCAATGTTA 23
|||||
Db 82138 AAACAACTGCACCCCAATGTTA 82116
RESULT 33
AC112763 211635 bp DNA linear HTG 19-NOV-2002
LOCUS AC112763
DEFINITION Rattus norvegicus clone CH230-246P24, *** SEQUENCING IN PROGRESS

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

***, 7 unordered pieces.
AC112763
AC112763.4 GI:25072605
HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
Rattus norvegicus (Norway rat)
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Rattus.
1 (bases 1 to 211635)
Muzny,D.Marie., Metzker,M.Lee., Abranzon,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
Gebregorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
Guaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
Harvey,I., Haviak,P., Hawes,A., Henderson,N., Hernandez,J.,
Hernandez,R., Hines,S., Hladun,S.B., Hodgson,A., Hogues,W.,
Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
Lorensuhea,L., Louissegh,H., Lozado,R.J., Lu,X., Ma,J.,
Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Nair,L.,
Nwackemele,O., Okwuonu,G., Olarnpungoon,A., Pal,S., Parks,K.,
Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkuch,C.,
Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.,
Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
Sanders,W., Savary,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajls,D.,
Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Umani,K.,
Valas,R., Vera,V., Villanana,D., Waldron,L., Walker,B., Wang,J.,
Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
Williams,G., Willson,R., Wlarczyk,R., Woodson,H., Worley,K.,
Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
Weinstock,G. and Gibbs,R.A.
Direct Submission
Unpublished
2 (bases 1 to 211635)
Worley,K.C.
Direct Submission
Submitted (24-FEB-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 211635)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (19-NOV-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

COMMENT

On Nov 19, 2002 this sequence version replaced gi:23195195.
 The sequence in this assembly is a combination of BAC based reads
 and whole genome shotgun sequencing reads assembled using Atlas
 (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described
 in the feature table below represents a scaffold in the Atlas
 assembly (a 'contig-scaffold'). Within each contig-scaffold,
 individual sequence contigs are ordered and oriented, and separated
 by sized gaps filled with Ns to the estimated size. The sequence
 may extend beyond the ends of the clone and there may be sequence
 contigs within a contig-scaffold that consist entirely of whole
 genome shotgun sequence reads. Both end sequences and whole genome
 shotgun sequence only contigs will be indicated in the feature
 table.

----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: GRXJ
 Center clone name: CH230-246P24
 ----- Summary Statistics

Assembly program: Phrap; version 0.990329
 Consensus quality: 197380 bases at least Q40
 Consensus quality: 199965 bases at least Q30
 Consensus quality: 201900 bases at least Q20
 Estimated insert size: 199042; sum-of-contigs estimation
 Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

 * NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 7 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

* 1 197215: contig of 197215 bp in length
 * 197216 197315: gap of unknown length
 * 197316 201157: contig of 3842 bp in length
 * 201158 201257: gap of unknown length
 * 201258 204646: contig of 3389 bp in length
 * 204647 204746: gap of unknown length
 * 204747 205898: contig of 1152 bp in length
 * 205899 207611: gap of unknown length
 * 207612 207711: contig of 1613 bp in length
 * 207712 209032: contig of 1321 bp in length
 * 209033 209132: gap of unknown length
 * 209133 211635: contig of 2503 bp in length.

FEATURES
 source : 1. 211635
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-246P24"
 1. 2887
 /note="wgs end extension
 clone_end:Sp6"
 2938. 4634
 /note="wgs end extension
 clone_end:Sp6"
 5539. 7053
 /note="wgs end extension
 clone_end:Sp6"
 complement(8698. 9583)
 /note="clone boundary
 clone_end:Sp6
 site:
 end sequence:BZ168459"
 189679. 190526

misc_feature

misc_feature

misc_feature

misc_feature

/note="clone boundary
 clone_end:T7
 site:
 end sequence:BZ168458"
 191302. 193241
 /note="wgs end extension
 clone_end:T7"
 194084. 195139
 /note="wgs end extension
 clone_end:T7"
 195560. 197215
 /note="wgs end extension
 clone_end:T7"
 197216. 197315
 /estimated length=unknown
 197316. 198590
 /note="wgs end extension
 clone_end:T7"
 201158. 201257
 /estimated length=unknown
 201258. 202440
 /note="wgs end extension
 clone_end:T7"
 204647. 204746
 /estimated length=unknown
 205899. 205998
 /estimated length=unknown
 207612. 207711
 /estimated length=unknown
 209033. 209132
 /estimated length=unknown

ORIGIN

Query Match 79.2%; Score 19.8; DB 14; Length 211635;
 Best Local Similarity 91.3%; Pred. No. 1.6e+02;
 Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 AAAAAGCTGCATCCCCCAATGTTA 23
 |||||
 Db 116307 AAAAAGCTGCATCCCCCACTTTA 116329

RESULT 34

BC045302/c 2124 bp mRNA linear VRT 08-MAR-2005
 LOCUS
 DEFINITION
 Danio rerio selenophosphate synthetase 1, mRNA (cdna clone
 MGC:55304 IMAGE:2600344), complete cds.

BC045302

ACCESSION

VERSION

BC045302.1 GI:28279506

KEYWORDS

SOURCE

ORGANISM

Danio rerio (zebrafish)

Danio rerio

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Actinopterygii; Neopterygii; Teleostei; Ostariophysi;

Cypriniformes; Cyprinidae; Danio.

1 (bases 1 to 2124)

Strausberg RL, Feingold EA, Grouse LH, Derge JG, Klausner RD,

Collins FS, Wagner L, Shenmen CM, Schuler GD, Altschul SF, Zeeberg

B, Buetow KH, Schaefer CF, Bhat NK, Hopkins RF, Jordan H, Moore T,

Max SI, Wang J, Haieff F, Diatchenko L, Marusina K, Farmer AA, Rubin

GM, Hong L, Stapleton M, Soares MB, Bonaudo MF, Casavant TL,

Scheetz TE, Brownstein MJ, Ustin TB, Toshiyuki S, Carninci P,

Prange C, Raha SS, Loquellano NA, Peters GJ, Abramson RD, Mulliahy

SJ, Bosak SA, McSwan PJ, McKernan KJ, Malek JA, Gunaratne PH,

Richards S, Worley KC, Hale S, Garcia AM, Gay LJ, Hulyk SW,

Villalón DK, Muzny DM, Sodergren EJ, Lu X, Gibbs RA, Fahey J,

Helton E, Kettelman M, Madan A, Rodrigues S, Sanchez A, Whiting M,

Madan A, Young AC, Shevchenko Y, Bouffard GG, Blakesley RW,

Touchman JW, Green ED, Dickson MC, Rodriguez AC, Grimwood J,

Schmutz J, Myers RM, Butterfield YS, Krzywinski MI, Skalska U,

Smailus DE, Schnerch A, Schein JE, Jones SJ and Marra MA.

CONSRTM

TITLE

Generation and initial analysis of more than 15,000 full-length

JOURNAL	human and mouse cDNA sequences	LOCUS	10-719-900-25
PURNED	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)	DEFINITION	12477932
REFERENCE	2 (bases 1 to 2124)	ACCESSION	BX276107
AUTHORS	Director MGC Project.	VERSION	BX276107.12
TITLE	Direct Submission	GI	50300026
JOURNAL	Submitted (31-JAN-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 1A03, Bethesda, MD 20892-2590, USA	HTG	
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov	SOURCE	Danio rerio (zebrafish)
COMMENT	Contact: MGC help desk Email: cgapbs-r@mail.nih.gov Tissue Procurement: Dr. Sumio Sugano and Dr. Ko-ichi Kawakami cDNA Library Preparation: Dr. Sumio Sugano and Dr. Ko-ichi Kawakami cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Institute for Systems Biology http://www.systemsbiology.org Contact: amadan@systemsbiology.org Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting	ORGANISM	Danio rerio (zebrafish)
FEATURES	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Series: IRAK Plate: 101 Row: b Column: 12 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 41054394. Location/Qualifiers	REFERENCE	1 (bases 1 to 46296)
source	1. .2124 /organism="Danio rerio" /mol_type="mRNA" /strain="AB" /db_xref="taxon:7955" /clone="MGC:55304 IMAGE:2600344" /tissue_type="Whole body, adult, (one male and one female, including unfertilized eggs)" /clone_lib="Sugano Kawakami zebrafish DRB" /lab_host="DH10B" /notes="Vector: pME18S-FL3"	AUTHORS	Gray, E.
gene	1. .2124 /genes="sephs1" /notes="synonym: fc49b09" /db_xref="GeneID:324947" /db_xref="ZFIN:ZDB-GENE-030131-3670" 115. .1293 /genes="sephs1" /codon_start=1 /product="selenophosphate synthetase" /protein_id="AAH45302.1" /db_xref="GI:28279507" /db_xref="GeneID:324947" /db_xref="ZFIN:ZDB-GENE-030131-3670"	TITLE	Submitted (14-JUL-2004) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: zfsh-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk On Jul 14, 2004 this sequence version replaced gi:48843323.
CDS	translation="MSVRESFPNPSYELDKNFRLTRFTELKGTGCKVPQDVLQKLLS LOENHQEDQFLGAVPRLGIGMDTCVPLRHGGLSLVQTTDYIPIVDDPYMMGRI ACANVLSDLYAMGVTCDDNMLLGVSNKLTGERKGVMLPLVIOGFKDASEAGTSVT GCGVTINPVLGGVATVCCQNEFIMPDPNAVPGDVLVLTKEPLGTQVAVHQLDIP EKWNKIKLVVTDVELATQEARMLNARLNRTAAGLMHTFNAAHATIDITFGILGHAQ NLARQQRVTSFVHNLPLVKAAVSKACGNMFLMHGTCPTSGGLLICILIPRQAA RFCAETKSPKYGEGHQAWIIGIVEKGNRTARIIDKPRIIEVAPQVATQNVNTTPGATS"	JOURNAL	Direct Submission
ORIGIN	77.6%; Score 19.4; DB 5; Length 2124; Best Local Similarity 95.2%; Pred. No. 3.4e+02; Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	COMMENT	During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above. This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Swi: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep Clone-derived Zebrafish pUC subclones occasionally display inconsistency over the length of mononucleotide A/T runs and conserved TA repeats. Where this is found the longest good quality representation will be submitted. Repeat names beginning 'Dr' were identified by the Recon repeat discovery system (Zhirong Bao and Sean Eddy, submitted), and those beginning 'drr' were identified by Rick Waterman (Stephen Johnson lab, WashU). For further information see http://www.sanger.ac.uk/Projects/D_rerio/fishmask.shtml CH211-218C17 is from a CHORI-211 BAC library VECTOR: pTABAC2.1.
Query Match	77.6%; Score 19.4; DB 5; Length 46296; Best Local Similarity 95.2%; Pred. No. 2.7e+02; Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	FEATURES	Location/Qualifiers
Qy	2 AAAAAGTCATCCCAATGTT 22	1. 46296 /organism="Danio rerio" /mol_type="genomic DNA" /db_xref="taxon:7955" /clone="CH211-218C17" /clone_lib="CHORI-211"	source
Db	1910 AGAAACTGCATCCCAATGTT 1890	ORIGIN	
RESULT 35			
BX276107/c			

misc_feature
75582..85230
/notes="assembly fragment:00174
fragment_chain:1"
85331..91591
/notes="assembly fragment:00023
fragment_chain:1"
91692..113878
/notes="assembly fragment:00773
fragment_chain:1
clone_end:r7
vector side:right"
113975..135260
/notes="assembly fragment:00469
fragment_chain:2"
135361..140962
/notes="assembly fragment:00089
fragment_chain:2"

ORIGIN

Query Match 77.6%; Score 19.4; DB 14; Length 140962;
Best Local Similarity 95.2%; Pred. No. 2.5e+02;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AAAAACTGCATCCCAATGT 21
|||||
Db 86821 AAAAACTGCATCCCAATGT 86841

RESULT 38

AL606513

LOCUS

DEFINITION

AL606513 172243 bp DNA linear ROD 17-DEC-2004
Mouse DNA sequence from clone Rp23-277B9 on chromosome 13 Contains
the the Viri5, Viri14, Viri10, Viri2, Virh5 and Virh6 genes for
vomeronasal 1 receptor 15, I14, I10, I2, H5 and H6, two Virh and
two Viri pseudogenes and an ATP-binding cassette sub-family F
(GCN20) member 2 (Abcf2) pseudogene, complete sequence.

ACCESSION

AL606513.3 GI:118476767

VERSION

HTG; Abcf2; Virh5; Virh6; Viri10; Viri14; Viri2; Viri5.

KEYWORDS

Mus musculus (house mouse)

SOURCE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 172243)
Lovell,J.

REFERENCE

Direct Submission

AUTHORS

Submitted (16-DEC-2004) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vega@sanger.ac.uk
Clone requests: Clonerequest@sanger.ac.uk

TITLE

On Feb 1, 2002 this sequence version replaced gi:18181711.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep -----

JOURNAL

Genome Center
Center: UK Medical Research Council

Center code: UK-MRC

Web site: http://mrcseq.har.mrc.ac.uk

Contact: mouse@har.mrc.ac.uk

COMMENT

This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one subclone; and the assembly was confirmed by restriction digest,
except on the rare occasion of the clone being a YAC.
Sequence from the Mouse Genome Sequencing Consortium whole genome
shotgun may have been used to confirm this sequence. Sequence data
from the whole genome shotgun alone has only been used where it has

a phred quality of at least 30.
Rp23-277B9 is from the RPCI-23 Mouse BAC Library
constructed by the group of Pieter de Jong.
For further details see http://www.chori.org/bacpac/home.htm
VECTOR: pBAC3.6.

FEATURES

source

Location/Qualifiers

1..172243
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="13"
/clone="RP23-277B9"
/clone_lib="RPCI-23"
complement(20246..21142)
/gene="Vlr15"
/locus_tag="RP23-277B9.7-001"
complement(20246..21142)
/gene="Vlr15"
/locus_tag="RP23-277B9.7-001"
complement(20246..21142)
/gene="Vlr15"
/locus_tag="RP23-277B9.7-001"
complement(20246..21142)
/locus_tag="RP23-277B9.7-001"
complement(20246..21142)
/standard_name="OTTMUSP00000000510"
/note="match: proteins: Q8R261"
/codon_start=1
/protein_id="CAI24853.1"
/db_xref="GI:56205067"
/translation="MNKITQAIVLSLAGPGVGNILVPMRYVYTSALGTEKRPIDLI
LIHLASNLIIICSTGVTDIVTFVFRNFDGIDGCKMVLARWARGLSICTCLLSV
VQAVTISPTTIWTKPQSCQVLFPFLLEFWINVLISNLLSYIKAGSSINRSVAA
TFIGHCVMLPSRHIIKWLFLSLMTLRDVIPOSLSGSSGNALHYKHKHVLVLRSS
FAGNSPEIRATWSVLILMAFLFFYVWDEILSFYTGFTVTHDSTLLNIQKVLFGY
ASFSPVILSRDHPVPPNVLHAH"
complement(31384..33122)
/locus_tag="RP23-277B9.2-001"
/pseudo
complement(31384..33122)
/locus_tag="RP23-277B9.2-001"
/note="match: proteins: Q8R270 Q8R271 Q8R272 Q8R280 Q8R281
Q8R274 Q8R275 Q8R276 Q8R268 Q8R279"
/pseudo
/codon_start=1
42897..43793
/gene="Vlrh14"
/locus_tag="RP23-277B9.11-001"
42897..43793
/gene="Vlrh14"
/locus_tag="RP23-277B9.11-001"
42897..43793
/gene="Vlrh14"
/locus_tag="RP23-277B9.11-001"
/standard_name="OTTMUSP00000000512"
/note="match: proteins: Q8R268 Q8R278 Q8R270 Q8R280
Q8R272"
/codon_start=1
/protein_id="CAI24855.1"
/db_xref="GI:56205068"
/translation="MVLOFIKETIPLEFMTVMGILGNMVSVHYMWSWGSPKPKIHL
ILTLSPNTNIIILAKGKQKTIYVGLRNFDDIGCKIIVYLRVARGLSICTSLLT
VQAIISPRASGWSRLPKGAWHILPFFSFWILNGLISNLIHSITSTGLNMSQLN
NSKNYCYFMPESREIKWIVLPLMVLURDAVFOGAMGASGHMIFLLKHKHVLVLRSS
KLLVTPPELRAOSVLLMLCFYFFYWTDCALSLFLSLGDSLSMINIQKVLTLGY
AVFSPVLVLRDGLLPACWHAQ"
complement(52854..53331)
/locus_tag="RP23-277B9.5-001"
/pseudo
complement(52854..53331)
/locus_tag="RP23-277B9.5-001"
/note="match: proteins: Q8R265 Q8R266 Q8R274 Q8R275
Q8R276 Q8R277 Q8R279 Q8R272"
/pseudo
/codon_start=1
complement(53705..54173)


```
* 94584 185950: contig of 91367 bp in length.
Location/Qualifiers
source
1..185950
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="1"
/clone="RP11-321D23"
/clone_lib="RPC1-11.2"
1..69804
/note="assembly_fragment:01327
clone_end:SP6
vector_side:left"
misc_feature
69905..86631
/note="assembly_fragment:03228
fragment_chain:1"
misc_feature
86732..94483
/note="assembly_fragment:03627
fragment_chain:1"
misc_feature
94584..185950
/note="assembly_fragment:01854
fragment_chain:1
clone_end:T7
vector_side:right"

ORIGIN
Query Match 77.6%; Score 19.4; DB 14; Length 185950;
Best Local Similarity 95.2%; Pred. No. 2.5e+02;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AAAAAGTCATCCCAATGTT 22
|||||
Db 95613 AAAAAGTCATCCCAATATT 95593

RESULT 40
CR936412/c
LOCUS
DEFINITION
228881 bp DNA linear HTG 29-JUL-2005
Danio rerio chromosome 18 clone DKEY-181M21, WORKING DRAFT
SEQUENCE.
CR936412
CR936412.10 GI:71142827
HTG; HTGS_PHASE2; HTGS_ACTIVEFIN; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE
Danio rerio (zebrafish)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 228881)
Leongamornlert,D.
Direct Submission
Submitted (28-JUL-2005) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
zfsh-help@sanger.ac.uk
zfsh-help@sanger.ac.uk Clone requests:
http://www.sanger.ac.uk/Projects/D_rerio/fags.shtml#dataeight
On Jul 25, 2005 this sequence version replaced gi:66392990.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: zfsh-help@sanger.ac.uk
----- Project Information
Center project name: zK181M21
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; 100% of reads
Consensus quality: 227384 bases at least Q40
Consensus quality: 227452 bases at least Q30
Consensus quality: 227477 bases at least Q20
Insert size: 228881; sum-of-contigs
Insert size: 214479; 21.4% error; agarose-fp
Quality coverage: 9.13x in Q20 bases; sum-of-contigs Quality
coverage: 9.79x in Q20 bases; agarose-fp

----- NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 228881: contig of 228881 bp in length.
Location/Qualifiers
source
1..228881
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:9555"
/chromosome="18"
/clone="DKEY-181M21"
/clone_lib="DanioKey"
1..228881
/note="assembly_fragment:01220"

FEATURES
misc_feature
1..228881
/note="assembly_fragment:01220"

ORIGIN
Query Match 77.6%; Score 19.4; DB 14; Length 228881;
Best Local Similarity 95.2%; Pred. No. 2.4e+02;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AAAAAGTCATCCCAATGTT 22
|||||
Db 37197 AGAAAGTCATCCCAATGTT 37177

RESULT 41
AY598742
LOCUS
DEFINITION
479 bp DNA linear INV 06-JUL-2005
Trilobodrilus axi cytochrome c oxidase subunit I gene, partial cds;
mitochondrial.
ACCESSION
AY598742
VERSION
AY598742.1 GI:51317447
KEYWORDS
SOURCE
mitochondrion Trilobodrilus axi
ORGANISM
Trilobodrilus axi
Eukaryota; Metazoa; Annelida; Polychaeta; Scoleleida; Dinophillida;
Dinophillidae; Trilobodrilus.
1 (bases 1 to 479)
Struck,T.H. and Purschke,G.
The sister group relationship of Aeolosomatidae and Potamodrilidae
(Annelida, 'Polychaeta') - a molecular phylogenetic approach based
on 18S rDNA and Cytochrome Oxidase I
Zool. Anz. 243, 281-293 (2005)
2 (bases 1 to 479)
Struck,T.H. and Purschke,G.
Direct Submission
Submitted (13-APR-2004) Life Science Department, Auburn University,
College of Mathematics and Science, 101 Rouse Building, Auburn, AL
36849, USA
Location/Qualifiers
1..479
/organism="Trilobodrilus axi"
/organelle="mitochondrion"
/mol_type="genomic DNA"
/db_xref="taxon:188474"
/country="Germany: List, North Sea island of Sylt"
<1..>479
/codon_start=1
/transl_table=5
/product="cytochrome c oxidase subunit I"
/protein_id="AAT99901.1"
/db_xref="GI:51317448"
/translation="VWAXXMFVGMVDVTRATMTIAVPTGKVFSLATINGS
PKMTEAPMLWAVGFIFLFTMGTLTGILLANSIDIVLHDTYVVAHFHVLVSMGAVFA
MFAGETHWFPPLFTGVNMHPRWTKAHFYMMFLGNLTFFPQHFGLGLSGMPRRYSYDPA
"

FEATURES
source
1..479
/organism="Trilobodrilus axi"
/organelle="mitochondrion"
/mol_type="genomic DNA"
/db_xref="taxon:188474"
/country="Germany: List, North Sea island of Sylt"
<1..>479
/codon_start=1
/transl_table=5
/product="cytochrome c oxidase subunit I"
/protein_id="AAT99901.1"
/db_xref="GI:51317448"
/translation="VWAXXMFVGMVDVTRATMTIAVPTGKVFSLATINGS
PKMTEAPMLWAVGFIFLFTMGTLTGILLANSIDIVLHDTYVVAHFHVLVSMGAVFA
MFAGETHWFPPLFTGVNMHPRWTKAHFYMMFLGNLTFFPQHFGLGLSGMPRRYSYDPA
"

CDS
```


to all pairs of overlapping reads to call NQS bases and SNPs. Alignments (between two reads) with less than 100 NQS bases or with SNP rate > 0.01 were discarded. To exclude alignment between two copies of a single read, comparisons between two reads that share 95% of their genome alignments (>=95% bases of read A and >=95% bases of read B were placed at the same locus of human genome) were discarded.

FEATURES

source
1..959
/organism="Pan troglodytes verus"
/mol_type="genomic DNA"
/sub_species="verus"
/db_xref="taxon:37012"
/clone_lib="Clint"
<1..959

STS

Query Match 76.8%; Score 19.2; DB 10; Length 959;
Best Local Similarity 87.5%; Pred. No. 4.5e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

ORIGIN

QY 2 AAAAACTGCATCCCAATGTTATG 25
|||||
DB 397 AAAAACTGCATATCCATTGTTATG 374

RESULT 44

BV543020
LOCUS BV543020 988 bp DNA linear STS 09-APR-2005
DEFINITION spi43b08.g1 Clint Pan troglodytes verus STS genomic, sequence tagged site.

ACCESSION

BV543020

VERSION

BV543020.1 GI:62434040

KEYWORDS

SOURCE

ORGANISM

Pan troglodytes verus

Pan troglodytes verus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Pan.
1 (bases 1 to 988)
Mikkelsen,T.S., Hillier,W.L., Eichler,E.E., Zody,M.C. and
Jaffe,D.B.
Initial Sequence of the Chimpanzee Genome and Comparison with the
Human Genome
Unpublished (2005)

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Contact: Michael C. Zody
Broad Institute of MIT and Harvard
320 Charles Street, Cambridge, MA 02141, USA
Tel: 6172580933
Fax: 6172580903
Email: mczody@broad.mit.edu
Primer A: No sequence submitted
Primer B: No sequence submitted
STS size: 988
Protocol:
23,021,928 chimpanzee whole genome shotgun reads were aligned to
the Human genome NCBI
Build 34 (hg16,July 2003). Chimp WGS reads were from 9 donors,
including Clint (Pan
troglodytes verus), 3 other Pan troglodytes verus chimps
(Donald,Karlien,Yvonne), 3 Pan
troglodytes troglodytes chimps (Noemie,Masuku,Clara) and 2 chimps
of unknown origin
(Gon,Unknown Chimp). Common names: Pan troglodytes verus is the
western chimp and Pan
troglodytes troglodytes is the central chimp. To be included in
chimpanzee SNP discovery, a
read must be at least 500bp in length, at least 50% of its base
calls must have Phred

score >= 20, at least 30% of its base calls must satisfy SNOS(30,25)(single strand NQS, the base in question has Phred score >= 30, the surrounding 10 bases in the read have Phred score >= 25), and the read must have at least 200 bp SNQS(30,25) bases. Reads not uniquely placed in the genome and read pairs whose two ends were not consistently placed were discarded. After above filtering, NQS(30,25) standard was applied to all pairs of overlapping reads to call NQS bases and SNPs. Alignments (between two reads) with less than 100 NQS bases or with SNP rate > 0.01 were discarded. To exclude alignment between two copies of a single read, comparisons between two reads that share 95% of their genome alignments (>=95% bases of read A and >=95% bases of read B were placed at the same locus of human genome) were discarded.

FEATURES

source
1..988
/organism="Pan troglodytes verus"
/mol_type="genomic DNA"
/sub_species="verus"
/db_xref="taxon:37012"
/clone_lib="Clint"
<1..988

STS

Query Match 76.8%; Score 19.2; DB 10; Length 988;
Best Local Similarity 87.5%; Pred. No. 4.4e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAACTGCATCCCAATGTTATG 25

|||||

DB 416 AAAAACTGCATATCCATTGTTATG 439

RESULT 45

AB122062/c
LOCUS AB122062 3079 bp mRNA linear INV 13-APR-2004
DEFINITION Crassostrea gigas HSP68 mRNA for 68kDa heat shock protein, complete cds.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

AB122062
Crassostrea gigas HSP68 mRNA for 68kDa heat shock protein, complete cds.
AB122062.1 GI:46359611
Crassostrea gigas (Pacific oyster)
Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Ostreoida;
Ostreidae; Ostreidae; Crassostrea.

REFERENCE

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

Yokoyama,Y.
Cloning and characterization of an inducible 68 kDa heat shock protein (HSP68) of the Japanese oyster (Crassostrea gigas)
Unpublished
2 (bases 1 to 3079)
Yokoyama,Y.
Direct Submission
Submitted (09-OCT-2003) Yoshihiro Yokoyama, Fukui Prefectural University, Department of Marine Bioscience, Gakuen-cyo 1-1, Obama, Fukui 917-0003, Japan (E-mail:yokoyama@pu.ac.jp, Tel:81770526300, Fax:81770526003)

FEATURES

source
1..3079
/organism="Crassostrea gigas"
/mol_type="mRNA"
/db_xref="taxon:29159"
/country="Japan"
1..3079
/gene="HSP68"
57..1898
/gene="HSP68"
/function="inducible chaperone"

gene

CDS

CDS	/gene="mrp-6"	
	/locus_tag="F20B6.3"	
misc_feature	complement(join(385..561,611..886,1271..886,1271..1519,1565..1768,2171..2274,2349..2469,2899..3005,3057..3207,3263..3342,3198..3666,3713..4219,5271..6124,6179..6329,7017..7142,7185..7563,7624..7826,8371..8603))	
	/gene="mrp-6"	
CDS	/locus_tag="F20B6.3"	
	/standard_names="F20B6.3"	
gene	/note="contains similarity to Pfam domains PF00864 (ABC transporter transmembrane region), PF00005 (ABC transporter); coded for by the following C. elegans cDNAs: yk297e4.5, yk326h11.5, yk605h4.5, yk679ds.5, yk297e4.3, yk605h4.3, yk679ds.3"	
	/codon_start=1	
misc_feature	/product="Multidrug resistance protein family protein 6"	
	/protein_id="AAA82317.2"	
CDS	/db_xref="GI:30794620"	
	/db_xref="WormBase:F20B6.3"	
gene	/translation="MIPHPQKTPDETKPDVNLESVEGVHPSEDSAGLISLTFSEFVG FYWTRNAQTDIDLEKPSGISAKYAAQKLESKWISGPKRTHFLVGVHCTKVSU MAHSLCLEELTRVAQPLLRLMELSYDIEEGFPYPWPAMQVALIISLSLFTVVCIL LFHYTHGLKGNHARVATVLLYKSLRISLTSFETNSGQIIQLLNTDAKLQKAF LFAHVWLCPLLMFFYAVITLMSFGFCLLGLFVMAIVSFQVYSYOLGVNRRNIGD KNDRLKVMSEILNGIKVIMKMGWENAPAEKIRILREDEIKAVGTSAFYSVMVMGIYF ISAKLGLFYIFACDIYFDEILITRGVFASSALFNSVRPLFPALFLPLGLLFGRELSTSRIDIFFLCLEESYRITLEDSTDDGKNEKEGUTLKKIGEKPSWTLKCDKAEKITSV FLNKVTAWYDEHDQKKKFNKALSDNDATVAGEKKDEPLHVYAVQDLSFEFV AGTCGIIIGSVGSGKSLILLITLSRMTGNLSVGFAYCAQBPWIPGSIQDNIV FCGSEKERYENALECLLITPDLOFANGDMTIVGDSGTLSGGQARVALARAVYAD ADIYLLDPLSAVDHVGKLVENIRGLFSLKVLVTHQIHFLSELDVILMKNNK MYGNSDELKSTFTDDFKLOAVETEVSQSDLFVDGPPSPPPSPVNIESMEKINKE SVIFQRANLSGSKMSETEYKAEIVTESRKSQITISWVTEYIRGVINPLKGIWILV CULITQLMNFDVWLNLKWLVDLMNADNLSKEFNRRNDVLGTFTNVTFQTYELT FIISTVMSILGVWSVPRNAQLMASRILHNSMDFCVLKTSEFFDKNSEGMILNRF SXDGLTDLLAFTGFVFGTFTFVGLVSMVRPIVILACAFITAGFYFCRKYV HRSRLKTRFAARSPLNTITSTVHGLSTIRAYKENEMIEKFCALHDVYVAAVNMG LLSARFNGICIDFLSVFSVIAIVVLQYETLTVGEVGLCLVCAVLQSLGFSEFWLRQ SAELENGWVSERVLEYTKLKEEERARRPHDAILDKLLNSWPTGEIEFENVSVMYQ ETKVLKNSVFLVQKQGVGRTGAGSTLKIIFGLKEYCSGRVTIDGVLDHLSL NFRRGMSIIPQEPVIFSGVTRENIDPFDVDDATWFALQCELRKSITKEKGLNVE LGQGVLSVSGOQLFCILARIMHRSKILVIDEATVNDVGTDACTQKTRIDFFKNAT IITVAHRLHTMNSDNVMVFEKGLVEFDKPIITLLDDQSSLFSRMAAQSGAENFDSLRAALAHAIIDQSSYDTS"	
	13248..13249	
misc_feature	/notes="Sli trans-splice acceptor; see yk759b01.5"	
	13252..14874	
gene	/gene="vha-12"	
	/locus_tag="F20B6.2"	
CDS	join(13252..13488,13538..13807,13859..14687,14735..14874)	
	/gene="vha-12"	
misc_feature	/standard_names="F20B6.2"	
	/note="contains similarity to Pfam domains PF02874 (ATP synthase alpha/beta family, beta-barrel domain), PF00006 (ATP synthase alpha/beta family, nucleotide-binding domain), PF00306 (ATP synthase alpha/beta chain, C terminal domain); coded for by the following C. elegans cDNAs: CB104433, OSTF003D11.1, OSTR003D11.1, yk86h1.5, yk102c11.3, yk102c11.5, yk213a8.3, yk213a8.5, yk214g9.5, yk239c12.3, yk243g10.5, yk291a2.5, yk304a10.3, yk304a10.5, yk337b9.5, yk339b7.5, yk367d3.5, yk372a8.3, yk372a8.5, yk386b10.3, yk386b10.5, yk448a3.5, yk448g12.5, yk453h4.5, yk464h1.5, yk49h10.5, yk498h3.3, yk498h3.5, yk536a4.5, yk628h3.3, yk628h3.5, yk633a3.5, yk640f5.5, yk647e12.5, yk652a10.3, yk652a10.5, yk668h12.5, yk669e11.5, yk759b01.5, yk755f10.5, yk777f09.3, yk777f09.5, yk778d04.5, yk784e03.3, yk784e03.5, yk785h09.3, yk785h09.5, yk786h08.3, yk786h08.5, yk790b12.3, yk790b12.5, yk799c12.3, yk799c12.5, yk800d03.3, yk800d03.5, yk836a11.3, yk837g06.3, yk837g06.5, yk840f01.3, yk840f01.5, yk848f01.3, yk848f01.5, yk850c10.3, yk852h06.3, yk"	

AP008209_048 4800001 4910000
AP008209_049 4900001 5010000
AP008209_050 5000001 5110000
AP008209_051 5100001 5210000
AP008209_052 5200001 5310000
AP008209_053 5300001 5410000
AP008209_054 5400001 5510000
AP008209_055 5500001 5610000
AP008209_056 5600001 5710000
AP008209_057 5700001 5810000
AP008209_058 5800001 5910000
AP008209_059 5900001 6010000
AP008209_060 6000001 6110000
AP008209_061 6100001 6210000
AP008209_062 6200001 6310000
AP008209_063 6300001 6410000
AP008209_064 6400001 6510000
AP008209_065 6500001 6610000
AP008209_066 6600001 6710000
AP008209_067 6700001 6810000
AP008209_068 6800001 6910000
AP008209_069 6900001 7010000
AP008209_070 7000001 7110000
AP008209_071 7100001 7210000
AP008209_072 7200001 7310000
AP008209_073 7300001 7410000
AP008209_074 7400001 7510000
AP008209_075 7500001 7610000
AP008209_076 7600001 7710000
AP008209_077 7700001 7810000
AP008209_078 7800001 7910000
AP008209_079 7900001 8010000
AP008209_080 8000001 8110000
AP008209_081 8100001 8210000
AP008209_082 8200001 8310000
AP008209_083 8300001 8410000
AP008209_084 8400001 8510000
AP008209_085 8500001 8610000
AP008209_086 8600001 8710000
AP008209_087 8700001 8810000
AP008209_088 8800001 8910000
AP008209_089 8900001 9010000
AP008209_090 9000001 9110000
AP008209_091 9100001 9210000
AP008209_092 9200001 9310000
AP008209_093 9300001 9410000
AP008209_094 9400001 9510000
AP008209_095 9500001 9610000
AP008209_096 9600001 9710000
AP008209_097 9700001 9810000
AP008209_098 9800001 9910000
AP008209_099 9900001 10010000
AP008209_100 10000001 10110000
AP008209_101 10100001 10210000
AP008209_102 10200001 10310000
AP008209_103 10300001 10410000
AP008209_104 10400001 10510000
AP008209_105 10500001 10610000
AP008209_106 10600001 10710000
AP008209_107 10700001 10810000
AP008209_108 10800001 10910000
AP008209_109 10900001 11010000
AP008209_110 11000001 11110000
AP008209_111 11100001 11210000
AP008209_112 11200001 11310000
AP008209_113 11300001 11410000
AP008209_114 11400001 11510000
AP008209_115 11500001 11610000
AP008209_116 11600001 11710000
AP008209_117 11700001 11810000
AP008209_118 11800001 11910000
AP008209_119 11900001 12010000
AP008209_120 12000001 12110000

AP008209_121 12100001 12210000
AP008209_122 12200001 12310000
AP008209_123 12300001 12410000
AP008209_124 12400001 12510000
AP008209_125 12500001 12610000
AP008209_126 12600001 12710000
AP008209_127 12700001 12810000
AP008209_128 12800001 12910000
AP008209_129 12900001 13010000
AP008209_130 13000001 13110000
AP008209_131 13100001 13210000
AP008209_132 13200001 13310000
AP008209_133 13300001 13410000
AP008209_134 13400001 13510000
AP008209_135 13500001 13610000
AP008209_136 13600001 13710000
AP008209_137 13700001 13810000
AP008209_138 13800001 13910000
AP008209_139 13900001 14010000
AP008209_140 14000001 14110000
AP008209_141 14100001 14210000
AP008209_142 14200001 14310000
AP008209_143 14300001 14410000
AP008209_144 14400001 14510000
AP008209_145 14500001 14610000
AP008209_146 14600001 14710000
AP008209_147 14700001 14810000
AP008209_148 14800001 14910000
AP008209_149 14900001 15010000
AP008209_150 15000001 15110000
AP008209_151 15100001 15210000
AP008209_152 15200001 15310000
AP008209_153 15300001 15410000
AP008209_154 15400001 15510000
AP008209_155 15500001 15610000
AP008209_156 15600001 15710000
AP008209_157 15700001 15810000
AP008209_158 15800001 15910000
AP008209_159 15900001 16010000
AP008209_160 16000001 16110000
AP008209_161 16100001 16210000
AP008209_162 16200001 16310000
AP008209_163 16300001 16410000
AP008209_164 16400001 16510000
AP008209_165 16500001 16610000
AP008209_166 16600001 16710000
AP008209_167 16700001 16810000
AP008209_168 16800001 16910000
AP008209_169 16900001 17010000
AP008209_170 17000001 17110000
AP008209_171 17100001 17210000
AP008209_172 17200001 17310000
AP008209_173 17300001 17410000
AP008209_174 17400001 17510000
AP008209_175 17500001 17610000
AP008209_176 17600001 17710000
AP008209_177 17700001 17810000
AP008209_178 17800001 17910000
AP008209_179 17900001 18010000
AP008209_180 18000001 18110000
AP008209_181 18100001 18210000
AP008209_182 18200001 18310000
AP008209_183 18300001 18410000
AP008209_184 18400001 18510000
AP008209_185 18500001 18610000
AP008209_186 18600001 18710000
AP008209_187 18700001 18810000
AP008209_188 18800001 18910000
AP008209_189 18900001 19010000
AP008209_190 19000001 19110000
AP008209_191 19100001 19210000
AP008209_192 19200001 19310000
AP008209_193 19300001 19410000

Consensus quality: 157076 bases at least Q40
Consensus quality: 160029 bases at least Q30
Consensus quality: 161054 bases at least Q20
Insert size: 158000; agarose-fp
Quality coverage: 4.9 in Q20 bases; agarose-fp
Quality coverage: 4.7 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 11 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 193: contig of 193 bp in length
* 194 293: gap of 100 bp
* 294 1492: contig of 1199 bp in length
* 1493 1592: gap of 100 bp
* 1593 2731: contig of 1139 bp in length
* 2732 2831: gap of 100 bp
* 2832 6919: contig of 4088 bp in length
* 6920 7019: gap of 100 bp
* 7020 11262: contig of 4243 bp in length
* 11263 11362: gap of 100 bp
* 11363 23705: contig of 12343 bp in length
* 23706 23805: gap of 100 bp
* 23806 46137: contig of 22332 bp in length
* 46138 46237: gap of 100 bp
* 46238 66133: contig of 19896 bp in length
* 66134 66233: gap of 100 bp
* 66234 91720: contig of 25486 bp in length
* 91720 91819: gap of 100 bp
* 91820 119835: contig of 28016 bp in length
* 119836 119935: gap of 100 bp
* 119936 163328: contig of 43393 bp in length.

FEATURES

source

1. .163328
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="RP11-23N4"
/clone_lib="RP11-11 Human Male BAC"

misc_feature

1. .193
/note="assembly_fragment"
clone end:T7
vector side:right"

gap

194. .293
/estimated_length=100

misc_feature

294. .1492
/note="assembly_fragment"

gap

1493. .1592
/estimated_length=100

misc_feature

1593. .2731
/note="assembly_fragment"

gap

2732. .2831
/estimated_length=100

misc_feature

2832. .6919
/note="assembly_fragment"

gap

6920. .7019
/estimated_length=100

misc_feature

7020. .11262
/note="assembly_fragment"

gap

11263. .11362
/estimated_length=100

misc_feature

11363. .23705
/note="assembly_fragment"

gap

23706. .23805
/estimated_length=100

misc_feature
46238. .66133
/note="assembly_fragment"
66134. .66233
/estimated_length=100
66234. .91719
/note="assembly_fragment"
clone end:SP6
vector side:right"
91720. .91819
/estimated_length=100
91820. .119835
/note="assembly_fragment"
119836. .119935
/estimated_length=100
119936. .163328
/note="assembly_fragment"

ORIGIN

Query Match 76.8%; Score 19.2; DB 14; Length 163328;
Best Local Similarity 87.5%; Pred No. 3.1e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 2 AAAAACTGCATCCCAATGTTATG 25
|||||
Db 135995 AAAAACTGCATATCCAGTGTATG 136018
|||||

Search completed: February 3, 2006, 21:27:13

Job time : 883 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 3, 2006, 10:53:33 ; Search time 203.111 Seconds
(without alignments)
820.326 Million cell updates/sec

Title: US-10-719-900-25
Perfect score: 25
Sequence: 1 aaaaactgcattcccaatgttatg 25

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 150 summaries

Database : N Geneseq_21.*
1: Geneseqm1980s.*
2: Geneseqm1990s.*
3: Geneseqm2000s.*
4: Geneseqm2001as.*
5: Geneseqm2001bs.*
6: Geneseqm2002as.*
7: Geneseqm2002bs.*
8: Geneseqm2003as.*
9: Geneseqm2003bs.*
10: Geneseqm2003cs.*
11: Geneseqm2003ds.*
12: Geneseqm2004as.*
13: Geneseqm2004bs.*
14: Geneseqm2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	22.4	89.6	3977	10	AD86804
2	22.4	89.6	3977	11	AD31040
3	21.8	87.2	270	10	ACA56412
4	21.8	87.2	270	12	AD156208
5	21.8	87.2	491	5	ABAI1028
6	21.8	87.2	798	5	ABAI1400
7	21.8	87.2	798	5	ABAI15361
8	21.8	87.2	798	5	ABAI15360
9	21.8	87.2	798	5	ABAI15360
10	20.4	81.6	1914	8	ACA46717
11	20.4	81.6	1917	4	AAH53037
12	20.4	81.6	1953	6	ABN92144
13	20.4	81.6	1953	13	ADS01836
14	20.4	81.6	3276	4	AAH54876
15	19.8	79.2	95050	14	ADX98574
16	19.2	76.8	179651	10	ADL13813
17	18.8	75.2	53552	5	AAH13655
18	18.6	74.4	168174	6	ABT11173
19	18.6	74.4	168273	6	ABT11114

20	18.4	73.6	110000	14	ABE42401_27
21	18.2	72.8	521	10	AD857979
22	18.2	72.8	521	10	AD857983
23	18.2	72.8	521	10	AD860446
24	18.2	72.8	521	10	AD860727
25	18.2	72.8	521	10	AD860731
26	18.2	72.8	95683	11	ACN43954
27	18.2	72.8	152048	10	ADL13855
28	17.8	71.2	578	6	ABQ31618
29	17.8	71.2	578	6	ABQ31618
30	17.8	71.2	1083	8	ACA24290
31	17.8	71.2	110000	14	ABE3175_33
32	17.8	71.2	162147	14	ABE35716
33	17.6	70.4	461	9	ACH39621
34	17.6	70.4	549	4	AAH09944
35	17.6	70.4	596	4	AAH28698
36	17.6	70.4	596	4	ADG41894
37	17.6	70.4	596	11	AD197668
38	17.6	70.4	612	5	ABV56236
39	17.6	70.4	614	4	AAH28126
40	17.6	70.4	614	10	ADG41030
41	17.6	70.4	614	11	AD196804
42	17.6	70.4	1086	8	ACA53946
43	17.6	70.4	1506	8	ACA43061
44	17.6	70.4	1770	11	ACH94314
45	17.6	70.4	1803	11	ACH94414
46	17.6	70.4	1951	13	ADR60058
47	17.6	70.4	2975	10	ADB68941
48	17.6	70.4	3833	10	AD825037
49	17.6	70.4	4363	4	AAK68667
50	17.6	70.4	4963	4	AAH36538
51	17.6	70.4	4963	8	ABX59526
52	17.6	70.4	4963	12	ADJ30276
53	17.6	70.4	8423	5	ABA15397
54	17.6	70.4	10256	4	ABK42295
55	17.6	70.4	10256	9	ABQ60451
56	17.6	70.4	12563	6	ABQ79513
57	17.6	70.4	12563	12	ADQ19748
58	17.6	70.4	12563	13	ADQ89825
59	17.6	70.4	12563	14	AEA24024
60	17.6	70.4	12775	12	ADQ23893
61	17.6	70.4	13572	9	AAH58743
62	17.6	70.4	13819	5	ABA15398
63	17.6	70.4	15944	5	ABA15399
64	17.6	70.4	68200	14	ADX80722
65	17.6	70.4	110000	6	ABQ659245_06
66	17.6	70.4	110000	6	ABQ67197_05
67	17.6	70.4	110000	14	ABE39175_31
68	17.6	70.4	229301	14	ABE35719
69	17.4	69.6	2179	4	AAI58201
70	17.4	69.6	2179	5	ADQ98408
71	17.4	69.6	2179	9	ADBA48168
72	17.4	69.6	3708	6	ABA05449
73	17.4	69.6	3790	5	ADL46016
74	17.4	69.6	5546	10	ACC06073
75	17.4	69.6	5546	10	ADK64761
76	17.4	69.6	17528	6	ABL32601
77	17.2	68.8	427	13	AD860633
78	17.2	68.8	681	8	ACA23085
79	17.2	68.8	866	2	AAH90428
80	17.2	68.8	1012	9	AAH56875
81	17.2	68.8	1875	12	ADJ39903
82	17.2	68.8	3198	9	ACD39241
83	17.2	68.8	3198	9	ACD39240
84	17.2	68.8	3198	9	ACD39243
85	17.2	68.8	3198	9	ACD39242
86	17.2	68.8	3917	2	AAH31723
87	17.2	68.8	7422	4	ABL08910
88	17.2	68.8	7447	4	ABL19158
89	17.2	68.8	8132	4	ABL19158
90	17.2	68.8	36714	13	ABD33628
91	17.2	68.8	36724	14	ADZ13602
92	17.2	68.8	38605	11	ACN44050

Continuation (28 o
Ad857979 Human gen
Ad857983 Human gen
Ad860446 Human gen
Ad860727 Human gen
Ad860731 Human gen
Adn43954 Human gen
Adl13855 Osteoarth
Abq31618 Oligonuc
Aca24290 Prokaryot
Continuation (34 o
Aeb35716 L. pneumo
Ach39621 Human con
Aah09944 Human cde
Aas28698 Genomic s
Adg41894 Human res
Adl97668 Human res
Abv56236 Human pro
Aas28126 Novel cdn
Adg41030 Human res
Adi96804 Human res
Aca53946 Prokaryot
Aca43061 Prokaryot
Ach94314 Klebsiell
Ach94414 Klebsiell
Adr60058 Cotton cd
Adb68941 C. neofo
Ade25037 Plant gro
Aak68667 Human imm
Aal36538 Human mus
Abx59526 cDNA enco
Adj30276 Human mus
Abal5397 Human ner
Abk42295 Genomic s
Adb60451 Connectiv
Abq79513 TRRAP nuc
Adq19748 Human sof
Adq89825 Antagonis
Aea24024 Human PRO
Adg23893 Human sof
Ad85743 Human tra
Abal5398 Human ner
Abal5399 Human ner
Adx80722 Human man
Continuation (7 o
Continuation (6 of
Continuation (32 o
Aeb35719 L. pneumo
Aai58201 Human pol
Adq98408 DNA encod
Adb48168 Novel hum
Aba05449 Human tra
Adl46016 Human ova
Acc06073 Gene sequ
Adk64761 Disease t
Ab132601 Human imm
Adr60633 Cotton cd
Aca23085 Prokaryot
Aat90428 K-ras gen
Ad56875 Human C-t
Adj39903 Plant cdn
Acd39241 Mosquito
Acd39240 Mosquito
Acd39243 Mosquito
Acd39242 Mosquito
Aat31723 K-ras onc
Ab108910 Drosophil
Ab119158 Drosophil
Abd33628 Human can
Adz13602 Human can
Acn44050 Human gen

```
93 17.2 68.8 62231 11 ACN44176
94 17.2 68.8 86453 14 ADY25730
c 95 17.2 68.8 96588 9 ADA03026
c 96 17.2 68.8 96588 10 ADB72764
c 97 17.2 68.8 96588 10 ADC85506
c 98 17.2 68.8 96588 12 ADM74621
c 99 17.2 68.8 110000 2 AAX20248_04
c 100 17.2 68.8 110000 2 AAX20248_05
c 101 17.2 68.8 116277 2 AAX20249
c 102 17.2 68.8 177866 10 ADL13935
c 103 17.2 68.8 216215 10 ADF69167
c 104 17.2 68.8 264965 12 ADN16203
c 105 17.2 68.8 268685 6 ABS56563
c 106 17 68.0 134 3 AACL12915
c 107 17 68.0 171 3 AACL18716
c 108 17 68.0 243 10 ADH82265
c 109 17 68.0 317 13 ADU14854
c 110 17 68.0 426 14 ACL61932
c 111 17 68.0 495 10 ADK60905
c 112 17 68.0 529 5 ADL44781
c 113 17 68.0 613 5 ADL63641
c 114 17 68.0 654 10 ADF03192
c 115 17 68.0 660 14 AEA54313
c 116 17 68.0 697 13 ADQ54890
c 117 17 68.0 851 5 ADL38397
c 118 17 68.0 851 5 ADL73266
c 119 17 68.0 871 13 ADX28923
c 120 17 68.0 879 4 AA194139
c 121 17 68.0 918 9 ADA31381
c 122 17 68.0 1021 10 ADK58271
c 123 17 68.0 1173 13 ADX10353
c 124 17 68.0 1204 2 AAX13339
c 125 17 68.0 1204 6 ABS99134
c 126 17 68.0 1574 5 ABA16953
c 127 17 68.0 2000 13 ADX10967
c 128 17 68.0 2782 14 ADZ48893
c 129 17 68.0 3012 4 AAH98552
c 130 17 68.0 3089 4 ABL24022
c 131 17 68.0 3811 4 ABL24612
c 132 17 68.0 5582 14 ADX06207
c 133 17 68.0 5749 14 ADX06203
c 134 17 68.0 5759 14 ADX06205
c 135 17 68.0 5780 12 ADN03901
c 136 17 68.0 5780 13 ADP23421
c 137 17 68.0 5925 13 ADU05939
c 138 17 68.0 5927 11 ADN95930
c 139 17 68.0 5927 14 ADX06201
c 140 17 68.0 6916 4 AAK78177
c 141 17 68.0 9031 4 ABL24594
c 142 17 68.0 26636 4 ABL16650
c 143 17 68.0 35042 3 AAAB1454
c 144 17 68.0 45698 12 ADJ12490
c 145 17 68.0 107820 4 AAD16230
c 146 17 68.0 110000 3 AAAB1489_1
c 147 17 68.0 110000 6 ABQ74964_4
c 148 17 68.0 110000 6 ABQ74964_5
c 149 17 68.0 110000 14 AEA61169_0
c 150 17 68.0 130312 14 AEB39168
```

ALIGNMENTS

```
RESULT 1
ADE86804
ID ADE86804 standard; cDNA; 3977 BP.
XX
AC ADE86804;
XX
DT 29-JAN-2004 (first entry)
XX
DE Kv3.1 alpha-subunit gene.
XX
```

```
ss: gene; rat ; Kv3.1 ; alpha-subunit; voltage-gated potassium channel ;
KW GABAergic interneuron; brain; parvalbumin-containing interneuron;
KW repolarisation; fast-firing phenotype ; prefrontal cortex ;
KW schizophrenia; splice variant.
OS Rattus norvegicus.
XX
FH Key Location/Qualifiers
FT CDS 1162..2919
FT /*tag= a
FT /product= "Kv3.1 alpha subunit"
EP1348963-A1.
XX
PD 01-OCT-2003.
XX
PF 28-MAR-2002; 2002EP-000071114.
XX
PR 28-MAR-2002; 2002EP-000071114.
XX
XX (MITS-) MITSUBISHI PHARMA CORP.
Cochran S, Yamagami K, Ohashi Y;
WPI: 2003-805901/76.
P-PSDB; ADE86805.
XX
XX New polynucleotides and polypeptides, useful for diagnosing
schizophrenia, and for identifying a compound modulating the activity of
Kv3 channels.
XX
XX Claim 1; Page 9-16; 24pp; English.
XX
XX This sequence encodes the rat Kv3.1 alpha-subunit. Kv3.1 is a voltage-
gated potassium channel which is predominantly located within a subset of
GABAergic interneurons within the brain, namely the parvalbumin-
containing interneurons. Its function is to allow fast repolarisation of
the cells and thus, contributes to the fast-firing phenotype of these
neurons. Parvalbumin expression has been shown to be altered within the
prefrontal cortex in schizophrenia. There are two alternative splice
variants of the Kv3.1 alpha-subunit, Kv3.1a and Kv3.1b, with each variant
only differing at the intracellular C-terminal. The Kv3.1 alpha-subunit
polynucleotide and polypeptide are useful for diagnosing schizophrenia,
and for identifying compounds that modulate the expression of the
polypeptide or polynucleotide, or the activity of Kv3.1. They are also
useful for the identification of subjects who are predisposed to
schizophrenia, and for evaluating the efficacy of drugs for such
disorder, and monitoring the progress of patient symptoms involved in
clinical trials for the treatment of such disorder.
XX
XX Sequence 3977 BP; 825 A; 1268 C; 1113 G; 771 T; 0 U; 0 Other;
Query Match 89.6%; Score 22.4; DB 10; Length 3977;
Best Local Similarity 95.8%; Pred. No. 3;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 AAAAAGTGCATCCCAATGTTATG 25
| | | | | | | | | | | | | | | | | | | | |
Db 3915 AAAAAGTGCATCCCAATGTTATG 3938
RESULT 2
ADY31040
ID ADY31040 standard; DNA; 3977 BP.
XX
AC ADY31040;
XX
DT 05-MAY-2005 (first entry)
XX
DE Rat voltage-gated potassium channel Kv3.1 DNA.
XX
KW diagnosis; schizophrenia; potassium channel; gene; ds.
XX
```



```
OS Rattus norvegicus.
XX
XX Key Location/Qualifiers
XX CDS 1162..2919
XX /*tag= a
XX /*product= "voltage-gated potassium channel Kv3.1"
XX /*gene= "Kcnc1"
XX
XX EP1348964-A2.
XX
XX 01-OCT-2003.
XX
XX 27-MAR-2003; 2003EP-00007002.
XX
XX 28-MAR-2002; 2002EP-00007114.
XX
XX (MITS-) MITSUBISHI PHARMA CORP.
XX
XX Cochran S, Yamagami K, Ohashi Y;
XX
XX WPI; 2003-805902/76.
XX P-PSDB; ADY31041.
XX REFSEQ; NM_012856.
XX
XX New polynucleotides and polypeptides, useful for diagnosing
XX schizophrenia, and for identifying a compound modulating the activity of
XX Kv3 channels.
XX
XX Claim 1; SEQ ID NO 1; 50pp; English.
XX
XX The invention relates to the use of a polynucleotide comprising a
XX sequence of 3977, 2441 or 3410 base pairs, fully defined in the
XX specification, its fragment or complement, for diagnosis or treating
XX schizophrenia. The invention also claims a method for screening for a
XX compound regulating the expression of the polynucleotide comprising: (a)
XX bringing a test compound into contact with a cell capable of expressing
XX the polynucleotide, whose expression is desired to be controlled; (b)
XX detecting the expression of the polypeptide in the cell; and (c)
XX determining a compound promoting or suppressing the expression of the
XX polypeptide compared to a control. The polynucleotides and polypeptides
XX are useful for diagnosing schizophrenia, and for identifying a compound
XX modulating the activity of Kv3 channels. They are also useful for the
XX identification of subjects who are predisposed to schizophrenia, and for
XX evaluating the efficacy of drugs for such disorder, and monitoring the
XX progress of patient symptoms involved in clinical trials for the
XX treatment of such disorder. This sequence corresponds to the gene for rat
XX voltage-gated potassium channel Kv3.1.
XX
XX Sequence 3977 BP; 825 A; 1268 C; 1113 G; 771 T; 0 U; 0 Other;
XX
XX Query Match 89.6%; Score 22.4; DB 11; Length 3977;
XX Best Local Similarity 95.8%; Pred. No. 3;
XX Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 2 AAAAAGTCATCCCAATGTTATG 25
XX |||||
XX Db 3915 AAAAAGTCATCCCAATGTTATG 3938
XX
XX RESULT 3
XX ID ACA56412 standard; cDNA; 270 BP.
XX
XX ACA56412;
XX
XX 06-JUN-2003 (first entry)
XX
XX Norway rat signalling pathway polynucleotide probe SEQ ID NO 1010.
XX
XX Norway rat; probe; ss; array element; Parkinson's disease;
XX signalling pathway population; cancer; adenocarcinoma; leukaemia;
XX immunopathy; AIDS; asthma; neuropathy; Alzheimer's disease; microarray.
XX
```

```
OS Rattus norvegicus.
XX
XX US6500938-B1.
XX
XX 31-DEC-2002.
XX
XX 30-JAN-1998; 98US-00016434.
XX
XX 30-JAN-1998; 98US-00016434.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Au-Young J, Seilhamer JJ;
XX
XX WPI; 2003-352189/33.
XX
XX Combination of polynucleotide probes, useful as array elements in a
XX microarray for monitoring the expression of a number of target
XX polynucleotides.
XX
XX Claim 1; SEQ ID NO 1010; 65pp; English.
XX
XX The invention relates to a combination which, comprises a number of
XX polynucleotide probes comprising a sequence selected from one of the 1490
XX sequences mentioned in the specification. The combination is useful as an
XX array element in a microarray for monitoring the expression of a number
XX of target polynucleotides. The microarray is particularly useful in the
XX diagnosis and treatment of cancer and immunopathology and neuropathology.
XX The microarray is useful in diagnostics and treatment regimens, drug
XX discovery and development, toxicological and carcinogenicity studies,
XX forensics and pharmacogenomics. The microarray is also useful for
XX monitoring progression of diseases and for developing sophisticated
XX profiles for the effects of currently available therapeutic drugs. The
XX combination is also useful for purifying a subpopulation of mRNAs, CDNAS
XX and genomic fragments and in research and diagnostic applications. The
XX array can detect changes in expression in a large number of genes coding
XX for different signalling pathway populations which can be used to diagnose
XX various diseases including cancer e.g. adenocarcinoma and leukaemia,
XX immunopathies e.g. AIDS and asthma, neuropathies e.g. Alzheimer's disease
XX and Parkinson's disease. The present sequence represents a polynucleotide
XX probe of the invention. Note: The sequence data for this patent did not
XX form part of the printed specification but was obtained in electronic
XX format directly from USPTO at
XX seqdata.uspto.gov/sequence.html?DocID=06500938B1
XX
XX Sequence 270 BP; 87 A; 56 C; 52 G; 75 T; 0 U; 0 Other;
XX
XX Query Match 87.2%; Score 21.8; DB 10; Length 270;
XX Best Local Similarity 92.0%; Pred. No. 3.9;
XX Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
XX QY 1 AAAAAGTCATCCCAATGTTATG 25
XX |||||
XX Db 164 AAAAAGTCATCCCAATGTTATG 188
XX
XX RESULT 4
XX ID ADI56208 standard; DNA; 270 BP.
XX
XX ADI56208;
XX
XX 22-APR-2004 (first entry)
XX
XX Human polynucleotide probe #1010.
XX
XX Human; probe; ss; receptor-like polypeptide; transducing polypeptide;
XX effector-like polypeptide; cancer; immunopathology; neuropathology;
XX drug development; toxicology; carcinogenicity;
XX signalling pathway polypeptide; adrenal gland; bladder; bone;
XX bone marrow; brain; breast; cervix; tumour; immunopathology; AIDS;
XX diabetes; pancreatitis; osteoporosis; ulcerative colitis; neuropathology;
XX dementia; amnesia; epilepsy; Alzheimer's disease; depression.
XX
```

```
XX Homo sapiens.
OS
XX US2004010136-A1.
PN
XX
XX 15-JAN-2004.
PD
XX
XX 26-NOV-2002; 2002US-00305720.
XX
XX 30-JAN-1998; 98US-00016434.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Au-Young J, Seilhamer JJ;
PI
XX WPI; 2004-090520/09.
XX
XX New composition comprising polynucleotide probes, useful as array
PT elements in a microarray for monitoring the expression of target
PT polynucleotides or purifying a subpopulation of mRNAs, cDNA, or genomic
PT fragments.
XX
XX Claim 6; SEQ ID NO 1010; 73bp; English.
XX
XX The invention relates to a composition of polynucleotide probes
CC comprising first polynucleotide probes comprising at least a portion of a
CC gene encoding a receptor-like polypeptide, second polynucleotide probes
CC comprising at least a portion of a gene encoding a transducing
CC polypeptide and third polynucleotide probes comprising at least a portion
CC of a gene encoding an effector-like polypeptide. The probes of the
CC composition are useful as array elements in a microarray for monitoring
CC the expression of target polynucleotides. The microarray is useful in the
CC diagnosis and treatment of cancer, an immunopathology or a
CC neuropathology. It can also be used for drug discovery and development,
CC toxicological and carcinogenicity studies, forensics or pharmacogenomics.
CC Microarrays can also be used for monitoring the progression of diseases
CC that may be associated with the altered expression of signalling pathway
CC polypeptides. The composition can also be used to purify a subpopulation
CC of mRNAs, cDNAs, or genomic fragments in a sample. The expression profile
CC is also useful for the diagnosis and treatment of cancer, e.g. cancers of
CC the adrenal gland, bladder, bone, bone marrow, brain, breast or cervix,
CC an immunopathology, e.g. AIDS, diabetes, pancreatitis, osteoporosis or
CC ulcerative colitis, or a neuropathology, e.g. dementia, amnesia,
CC epilepsy, Alzheimer's disease or depression. This sequence represents a
CC human polynucleotide probe of the invention. Note: The sequence data for
CC this patent did not form part of the printed specification but was
CC obtained in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html.
XX
XX Sequence 270 BP; 87 A; 56 C; 52 G; 75 T; 0 U; 0 Other;
SQ
Query Match 87.2%; Score 21.8; DB 12; Length 270;
Best Local Similarity 92.0%; Pred. No. 3.9;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 AAAAACTGCATCCCAATGTTATG 25
Db 164 AAAAACTGCATCCCAATGTTATG 188
RESULT 5
ABAI1028
ID ABAI1028 standard; cDNA; 491 BP.
XX
XX ABAI1028;
XX
XX 23-JAN-2002 (first entry)
XX
XX Human nervous system related polynucleotide SEQ ID NO 35.
DE
DE Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;
XX immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;
XX antiparkinsonian; antisickling; antianaemic; antiarthritic; cancer;
XX
```

antirheumatic; hepatotropic; cerebrotropic; antiinflammatory;
antiallergic; antidiabetic; antiulcer; anticonvulsant; antifungal;
antiparasitic; cardiant; immune disorder; cardiovascular disorder;
neurological disease; infection; nephrotropic; gene therapy; vaccine; ss.
Homo sapiens.
WO200159063-A2.
16-AUG-2001.
17-JAN-2001; 2001WO-US001334.
31-JAN-2000; 2000US-0179065P.
04-FEB-2000; 2000US-0180628P.
24-FEB-2000; 2000US-0184664P.
02-MAR-2000; 2000US-0186350P.
16-MAR-2000; 2000US-0189874P.
17-MAR-2000; 2000US-0190076P.
18-APR-2000; 2000US-0198123P.
19-MAY-2000; 2000US-0205515P.
07-JUN-2000; 2000US-0209467P.
28-JUN-2000; 2000US-0214886P.
30-JUN-2000; 2000US-0215135P.
07-JUL-2000; 2000US-0216647P.
07-JUL-2000; 2000US-0216880P.
11-JUL-2000; 2000US-0217487P.
11-JUL-2000; 2000US-0217496P.
14-JUL-2000; 2000US-0218290P.
26-JUL-2000; 2000US-0220963P.
26-JUL-2000; 2000US-0220964P.
14-AUG-2000; 2000US-0224518P.
14-AUG-2000; 2000US-0224519P.
14-AUG-2000; 2000US-0225213P.
14-AUG-2000; 2000US-0225214P.
14-AUG-2000; 2000US-0225266P.
14-AUG-2000; 2000US-0225267P.
14-AUG-2000; 2000US-0225268P.
14-AUG-2000; 2000US-0225270P.
14-AUG-2000; 2000US-0225447P.
14-AUG-2000; 2000US-0225757P.
14-AUG-2000; 2000US-0225758P.
14-AUG-2000; 2000US-0225759P.
18-AUG-2000; 2000US-0226279P.
22-AUG-2000; 2000US-0226681P.
22-AUG-2000; 2000US-0226688P.
22-AUG-2000; 2000US-0227182P.
23-AUG-2000; 2000US-0227009P.
30-AUG-2000; 2000US-0228924P.
01-SEP-2000; 2000US-0229287P.
01-SEP-2000; 2000US-0229343P.
01-SEP-2000; 2000US-0229344P.
01-SEP-2000; 2000US-0229345P.
05-SEP-2000; 2000US-0229509P.
05-SEP-2000; 2000US-0229513P.
06-SEP-2000; 2000US-0230437P.
06-SEP-2000; 2000US-0230438P.
08-SEP-2000; 2000US-0231242P.
08-SEP-2000; 2000US-0231243P.
08-SEP-2000; 2000US-0231244P.
08-SEP-2000; 2000US-0231413P.
08-SEP-2000; 2000US-0231414P.
08-SEP-2000; 2000US-0232080P.
08-SEP-2000; 2000US-0232081P.
12-SEP-2000; 2000US-0231968P.
14-SEP-2000; 2000US-0232397P.
14-SEP-2000; 2000US-0232398P.
14-SEP-2000; 2000US-0232399P.
14-SEP-2000; 2000US-0232400P.
14-SEP-2000; 2000US-0232401P.
14-SEP-2000; 2000US-0233063P.
14-SEP-2000; 2000US-0233064P.
14-SEP-2000; 2000US-0233065P.

```
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 28-SEP-2000; 2000US-0236327P.
PR 28-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 20-OCT-2000; 2000US-0242221P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246529P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249246P.
PR 17-NOV-2000; 2000US-0249247P.
PR 17-NOV-2000; 2000US-0249256P.
PR 17-NOV-2000; 2000US-0249257P.
PR 17-NOV-2000; 2000US-0249259P.
PR 17-NOV-2000; 2000US-0249299P.
PR 01-DEC-2000; 2000US-0250391P.
PR 01-DEC-2000; 2000US-0251160P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.

PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-541565/60.
XX P-PSDB; ABB14702.
XX Nucleic acids encoding 3224 human nervous system antigen polypeptides,
XX useful for preventing, diagnosing and/or treating nervous system cancers
XX and metastases.
XX Claim 1; SEQ ID NO 35; 1701pp + Sequence Listing; English.
XX The invention relates to novel genes (AB11004-ABA1534) and proteins
XX (AB114678-AB18001) useful for preventing, treating or ameliorating
XX medical conditions e.g. by protein or gene therapy. The genes are
XX isolated from a range of human tissues disclosed in the specification.
XX The nucleic acids, proteins, antibodies and (ant)agonists are useful in
XX the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
XX ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
XX breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
XX disorders e.g. Addison's disease, allergies, diabetes mellitus, Crohn's disease,
XX anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
XX multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
XX cardiovascular disorders such as myocardial ischaemia; (d) wound healing
XX ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
XX infectious diseases such as viral, bacterial, fungal and parasitic
XX infections. Note: The sequence data for this patent did not form part of
XX the printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 491 BP; 137 A; 120 C; 100 G; 132 T; 0 U; 2 Other;
Query Match 87.2%; Score 21.8; DB 5; Length 491;
Best Local Similarity 92.0%; Pred. No. 4.2;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAACTGCATCCCAATGTTATG 25
Db 447 AAAAACTGCATCCCAATGTTATG 471

RESULT 6
ABA14400
ID ABA14400 standard; DNA; 798 BP.
XX ABA14400;
AC ABA14400;
XX
XX 23-JAN-2002 (first entry)
XX Human nervous system related polynucleotide SEQ ID NO 6731.
XX Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;
XX immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;
XX antiparkinsonian; antiskilling; antianaemic; antiarthritic; cancer;
XX antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
XX antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;
XX antiparasitic; cardiant; immune disorder; cardiovascular disorder;
XX neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.
XX Homo sapiens.
XX WO200159063-A2.
XX
XX 16-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US001334.
XX
XX 31-JAN-2000; 2000US-0179065P.
```

PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226688P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.

PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 20-OCT-2000; 2000US-0242221P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250391P.
PR 01-DEC-2000; 2000US-0251160P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-541565/60.

Nucleic acids encoding 3224 human nervous system antigen polypeptides,
useful for preventing, diagnosing and/or treating nervous system cancers
and metastases.

XX

PS Disclosure; SEQ ID NO 6731; 1701pp + Sequence Listing; English.

XX The invention relates to novel genes (ABAI1004-ABA21534) and proteins

CC (ABBI4678-ABBI8001) useful for preventing, treating or ameliorating

CC medical conditions e.g. by protein or gene therapy. The genes are

CC isolated from a range of human tissues disclosed in the specification.

CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in

CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and

CC ovarian cancer, and other cancers of the adrenal gland, bone, bone marrow,

CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune

CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic

CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,

CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)

CC cardiovascular disorders such as myocardial ischaemia; (d) wound healing

CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)

CC infectious diseases such as viral, bacterial, fungal and parasitic

CC infections. Note: The sequence data for this patent did not form part of

CC the printed specification, but was obtained in electronic format directly

CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX

SQ Sequence 798 BP; 228 A; 184 C; 173 G; 213 T; 0 U; 0 Other;

Query Match 87.2%; Score 21.8; DB 5; Length 798;

Best Local Similarity 92.0%; Pred. No. 4.5;

Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAACTGCATCCCAATGTTATG 25

Db 764 AAAAACTGCATCCCAATGTTATG 788

|||||

RESULT 7

ABAI5361

ID ABAI5361 standard; DNA; 798 BP.

XX

AC ABAI5361;

XX

DT 23-JAN-2002 (first entry)

XX

DE Human nervous system related polynucleotide SEQ ID NO 7692.

XX

KW Human; neurotropic; neuroprotective; cytostatic; dermatological; virucide;

KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;

KW antiparkinsonian; antiseizure; antianemic; antiarthritic; cancer;

KW antineuritic; hepatotropic; cerebroprotective; antiinflammatory;

KW antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;

KW antiparasitic; cardiac; immune disorder; cardiovascular disorder;

KW neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.

XX

OS Homo sapiens.

XX

XX WO200159063-A2.

XX

PD 16-AUG-2001.

XX

XX

PF 17-JAN-2001; 2001WO-US001334.

XX

XX 31-JAN-2000; 2000US-0179065P.

PR 04-FEB-2000; 2000US-0180628P.

PR 24-FEB-2000; 2000US-0184664P.

PR 02-MAR-2000; 2000US-0186350P.

PR 16-MAR-2000; 2000US-0189874P.

PR 17-MAR-2000; 2000US-0190076P.

PR 18-APR-2000; 2000US-0198123P.

PR 19-MAY-2000; 2000US-0205515P.

PR 07-JUN-2000; 2000US-0209467P.

PR 28-JUN-2000; 2000US-0214886P.

PR 30-JUN-2000; 2000US-0215135P.

PR 07-JUL-2000; 2000US-0216647P.

PR 07-JUL-2000; 2000US-0216880P.

PR 11-JUL-2000; 2000US-0217487P.

PR 11-JUL-2000; 2000US-0217496P.

PR 14-JUL-2000; 2000US-0218290P.

PR 26-JUL-2000; 2000US-0220963P.

PR 26-JUL-2000; 2000US-0220964P.

PR 14-AUG-2000; 2000US-0224518P.

PR 14-AUG-2000; 2000US-0224519P.

PR 14-AUG-2000; 2000US-0225213P.

PR 14-AUG-2000; 2000US-0225214P.

PR 14-AUG-2000; 2000US-0225266P.

PR 14-AUG-2000; 2000US-0225267P.

PR 14-AUG-2000; 2000US-0225268P.

PR 14-AUG-2000; 2000US-0225270P.

PR 14-AUG-2000; 2000US-0225447P.

PR 14-AUG-2000; 2000US-0225757P.

PR 14-AUG-2000; 2000US-0225758P.

PR 14-AUG-2000; 2000US-0225759P.

PR 18-AUG-2000; 2000US-0226279P.

PR 22-AUG-2000; 2000US-0226681P.

PR 22-AUG-2000; 2000US-0226868P.

PR 22-AUG-2000; 2000US-0227182P.

PR 23-AUG-2000; 2000US-0227009P.

PR 30-AUG-2000; 2000US-0228924P.

PR 01-SEP-2000; 2000US-0229287P.

PR 01-SEP-2000; 2000US-0229343P.

PR 01-SEP-2000; 2000US-0229344P.

PR 05-SEP-2000; 2000US-0229345P.

PR 05-SEP-2000; 2000US-0229509P.

PR 06-SEP-2000; 2000US-0229513P.

PR 06-SEP-2000; 2000US-0230437P.

PR 06-SEP-2000; 2000US-0230438P.

PR 08-SEP-2000; 2000US-0231242P.

PR 08-SEP-2000; 2000US-0231243P.

PR 08-SEP-2000; 2000US-0231244P.

PR 08-SEP-2000; 2000US-0231413P.

PR 08-SEP-2000; 2000US-0231414P.

PR 08-SEP-2000; 2000US-0232080P.

PR 08-SEP-2000; 2000US-0232081P.

PR 12-SEP-2000; 2000US-0231968P.

PR 14-SEP-2000; 2000US-0232397P.

PR 14-SEP-2000; 2000US-0232398P.

PR 14-SEP-2000; 2000US-0232399P.

PR 14-SEP-2000; 2000US-0232400P.

PR 14-SEP-2000; 2000US-0232401P.

PR 14-SEP-2000; 2000US-0233063P.

PR 14-SEP-2000; 2000US-0233064P.

PR 21-SEP-2000; 2000US-0234223P.

PR 21-SEP-2000; 2000US-0234274P.

PR 25-SEP-2000; 2000US-0234597P.

PR 26-SEP-2000; 2000US-0234998P.

PR 26-SEP-2000; 2000US-0235484P.

PR 27-SEP-2000; 2000US-0235834P.

PR 27-SEP-2000; 2000US-0235836P.

PR 29-SEP-2000; 2000US-0236327P.

PR 29-SEP-2000; 2000US-0236367P.

PR 29-SEP-2000; 2000US-0236368P.

PR 29-SEP-2000; 2000US-0236369P.

PR 29-SEP-2000; 2000US-0236370P.

PR 02-OCT-2000; 2000US-0236802P.

PR 02-OCT-2000; 2000US-0237037P.

PR 02-OCT-2000; 2000US-0237038P.

PR 02-OCT-2000; 2000US-0237039P.

PR 02-OCT-2000; 2000US-0237040P.

PR 13-OCT-2000; 2000US-0239935P.

PR 13-OCT-2000; 2000US-0239937P.

PR 20-OCT-2000; 2000US-0240960P.

PR 20-OCT-2000; 2000US-0241785P.

PR 20-OCT-2000; 2000US-0241786P.

PR 20-OCT-2000; 2000US-0241787P.

PR 20-OCT-2000; 2000US-0241808P.

PR 20-OCT-2000; 2000US-0241809P.

PR 20-OCT-2000; 2000US-0241826P.

PR 20-OCT-2000; 2000US-0242221P.

PR 01-NOV-2000; 2000US-0244617P.

PR 08-NOV-2000; 2000US-0246474P.

PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250319P.
PR 01-DEC-2000; 2000US-0251160P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 11-DEC-2000; 2000US-0251990P.
PR 05-JAN-2001; 2001US-0259678P.
(HUMA-) HUMAN GENOME SCI INC.
Rosen CA, Barash SC, Ruben SM;
WPI; 2001-541565/60.
Nucleic acids encoding 3224 human nervous system antigen polypeptides,
useful for preventing, diagnosing and/or treating nervous system cancers
and metastases.
Disclosure; SEQ ID NO 7692; 1701pp + Sequence Listing; English.
The invention relates to novel genes (ABAI1004-ABA21534) and proteins
(ABBI4678-ABBI8001) useful for preventing, treating or ameliorating
medical conditions e.g. by protein or gene therapy. The genes are
isolated from a range of human tissues disclosed in the specification.
The nucleic acids, proteins, antibodies and (ant)agonists are useful in
the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
ovarian cancer, and other cancers of the adrenal gland, bone, bone marrow,
breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
disorders e.g. Addison's disease, allergies, autoimmune haemolytic
anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
cardiovascular disorders such as myocardial ischaemia; (d) wound healing
; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)

CC infectious diseases such as viral, bacterial, fungal and parasitic
CC infections. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 798 BP; 227 A; 184 C; 174 G; 213 T; 0 U; 0 Other;

Query Match 87.2%; Score 21.8; DB 5; Length 798;

Best Local Similarity 92.0%; Pred. No. 4.5;

Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAGTCATGCCCAATGTTATG 25

||||| ||||| ||||| ||||| |||||
Db 764 AAAAAGTCATGCCCAATGTTATG 788

RESULT 8

ABAI5360

ID ABAI5360 standard; DNA; 798 BP.

XX ABAI5360;

XX 23-JAN-2002 (first entry)

DE Human nervous system related polynucleotide SEQ ID NO 7691.

KW Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;
KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnery;
KW antiparkinsonian; antiskilling; antianaemic; antitubercic; cancer;
KW antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
KW antiallergic; antidiabetic; antiulcer; anticonvulsant; antifungal;
KW antiparasitic; cardiant; immune disorder; cardiovascular disorder;
KW neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.

XX Homo sapiens.

OS WO200159063-A2.

PN 16-AUG-2001.

PD 17-JAN-2001; 2001WO-US001334.

PR 31-JAN-2000; 2000US-0179065P.

PR 04-FEB-2000; 2000US-0180628P.

PR 24-FEB-2000; 2000US-0184664P.

PR 02-MAR-2000; 2000US-0186350P.

PR 16-MAR-2000; 2000US-0189874P.

PR 17-MAR-2000; 2000US-0190076P.

PR 18-APR-2000; 2000US-0198123P.

PR 19-MAY-2000; 2000US-0205515P.

PR 07-JUN-2000; 2000US-0209467P.

PR 28-JUN-2000; 2000US-0214886P.

PR 30-JUN-2000; 2000US-0215135P.

PR 07-JUL-2000; 2000US-0216647P.

PR 07-JUL-2000; 2000US-0216880P.

PR 11-JUL-2000; 2000US-0217487P.

PR 11-JUL-2000; 2000US-0217496P.

PR 14-JUL-2000; 2000US-0218290P.

PR 26-JUL-2000; 2000US-0220963P.

PR 26-JUL-2000; 2000US-0220964P.

PR 14-AUG-2000; 2000US-0224518P.

PR 14-AUG-2000; 2000US-0224519P.

PR 14-AUG-2000; 2000US-0225213P.

PR 14-AUG-2000; 2000US-0225214P.

PR 14-AUG-2000; 2000US-0225266P.

PR 14-AUG-2000; 2000US-0225267P.

PR 14-AUG-2000; 2000US-0225268P.

PR 14-AUG-2000; 2000US-0225270P.

PR 14-AUG-2000; 2000US-0225447P.

PR 14-AUG-2000; 2000US-0225757P.

PR 14-AUG-2000; 2000US-0225758P.

PR 14-AUG-2000; 2000US-0225759P.

PR 18-AUG-2000; 2000US-0226279P.

PR 22-AUG-2000; 2000US-0226681P.
 PR 22-AUG-2000; 2000US-0226686P.
 PR 22-AUG-2000; 2000US-0227182P.
 PR 23-AUG-2000; 2000US-0227009P.
 PR 30-AUG-2000; 2000US-0228924P.
 PR 01-SEP-2000; 2000US-0229287P.
 PR 01-SEP-2000; 2000US-0229343P.
 PR 01-SEP-2000; 2000US-0229344P.
 PR 01-SEP-2000; 2000US-0229345P.
 PR 05-SEP-2000; 2000US-0229509P.
 PR 05-SEP-2000; 2000US-0229513P.
 PR 06-SEP-2000; 2000US-0230437P.
 PR 06-SEP-2000; 2000US-0230438P.
 PR 08-SEP-2000; 2000US-0231124P.
 PR 08-SEP-2000; 2000US-0231123P.
 PR 08-SEP-2000; 2000US-0231144P.
 PR 08-SEP-2000; 2000US-0231143P.
 PR 08-SEP-2000; 2000US-0231141P.
 PR 08-SEP-2000; 2000US-0232080P.
 PR 08-SEP-2000; 2000US-0232081P.
 PR 12-SEP-2000; 2000US-0231968P.
 PR 14-SEP-2000; 2000US-0232397P.
 PR 14-SEP-2000; 2000US-0232398P.
 PR 14-SEP-2000; 2000US-0232399P.
 PR 14-SEP-2000; 2000US-0232400P.
 PR 14-SEP-2000; 2000US-0232401P.
 PR 14-SEP-2000; 2000US-0233063P.
 PR 14-SEP-2000; 2000US-0233064P.
 PR 14-SEP-2000; 2000US-0233065P.
 PR 21-SEP-2000; 2000US-0234223P.
 PR 21-SEP-2000; 2000US-0234274P.
 PR 25-SEP-2000; 2000US-0234997P.
 PR 25-SEP-2000; 2000US-0234998P.
 PR 26-SEP-2000; 2000US-0235484P.
 PR 27-SEP-2000; 2000US-0235834P.
 PR 27-SEP-2000; 2000US-0235836P.
 PR 29-SEP-2000; 2000US-0236327P.
 PR 29-SEP-2000; 2000US-0236367P.
 PR 29-SEP-2000; 2000US-0236368P.
 PR 29-SEP-2000; 2000US-0236369P.
 PR 29-SEP-2000; 2000US-0236370P.
 PR 02-OCT-2000; 2000US-0236802P.
 PR 02-OCT-2000; 2000US-0237037P.
 PR 02-OCT-2000; 2000US-0237038P.
 PR 02-OCT-2000; 2000US-0237039P.
 PR 02-OCT-2000; 2000US-0237040P.
 PR 13-OCT-2000; 2000US-0239935P.
 PR 13-OCT-2000; 2000US-0239937P.
 PR 20-OCT-2000; 2000US-0240960P.
 PR 20-OCT-2000; 2000US-0241785P.
 PR 20-OCT-2000; 2000US-0241786P.
 PR 20-OCT-2000; 2000US-0241787P.
 PR 20-OCT-2000; 2000US-0241808P.
 PR 20-OCT-2000; 2000US-0241809P.
 PR 20-OCT-2000; 2000US-0241826P.
 PR 20-OCT-2000; 2000US-0242221P.
 PR 01-NOV-2000; 2000US-0244617P.
 PR 08-NOV-2000; 2000US-0246474P.
 PR 08-NOV-2000; 2000US-0246475P.
 PR 08-NOV-2000; 2000US-0246476P.
 PR 08-NOV-2000; 2000US-0246477P.
 PR 08-NOV-2000; 2000US-0246478P.
 PR 08-NOV-2000; 2000US-0246523P.
 PR 08-NOV-2000; 2000US-0246524P.
 PR 08-NOV-2000; 2000US-0246525P.
 PR 08-NOV-2000; 2000US-0246526P.
 PR 08-NOV-2000; 2000US-0246527P.
 PR 08-NOV-2000; 2000US-0246528P.
 PR 08-NOV-2000; 2000US-0246532P.
 PR 08-NOV-2000; 2000US-0246609P.
 PR 08-NOV-2000; 2000US-0246610P.
 PR 08-NOV-2000; 2000US-0246611P.
 PR 08-NOV-2000; 2000US-0246613P.

PR 17-NOV-2000; 2000US-0249207P.
 PR 17-NOV-2000; 2000US-0249208P.
 PR 17-NOV-2000; 2000US-0249209P.
 PR 17-NOV-2000; 2000US-0249210P.
 PR 17-NOV-2000; 2000US-0249211P.
 PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249244P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249264P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 01-DEC-2000; 2000US-0251160P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 05-DEC-2000; 2000US-0256719P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Barash SC, Ruben SM;
 XX WPI; 2001-541565/60.
 XX
 DR
 XX
 PT Nucleic acids encoding 3224 human nervous system antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating nervous system cancers
 PT and metastases.
 XX
 PS Disclosure; SEQ ID NO 7691; 1701pp + Sequence Listing; English.
 XX
 CC The invention relates to novel genes (ABA11004-ABA21534) and proteins
 CC (AB14678-AB18001) useful for preventing, treating or ameliorating
 CC medical conditions e.g. by protein or gene therapy. The genes are
 CC isolated from a range of human tissues disclosed in the specification.
 CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in
 CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
 CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
 CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
 CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
 CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
 CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
 CC cardiovascular disorders such as myocardial ischaemia; (d) wound healing
 CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
 CC infectious diseases such as viral, bacterial, fungal and parasitic
 CC infections. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 798 BP; 228 A; 184 C; 173 G; 213 T; 0 U; 0 Other;

Query Match 87.2%; Score 21.8; DB 5; Length 798;
 Best Local Similarity 92.0%; Pred. No. 4.5;
 Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 AAAAACTGCATCCCAATGTTATG 25
 |||||
 Db 764 AAAAACTGCATCCCAATGTTATG 788

PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 01-DEC-2000; 2000US-0251160P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 05-DEC-2000; 2000US-0256719P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 08-DEC-2000; 2000US-0251990P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Barash SC, Ruben SM;
 XX
 DR WPI; 2001-541565/60.
 XX
 PT Nucleic acids encoding 3224 human nervous system antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating nervous system cancers
 PT and metastases.
 XX
 PS Disclosure; SEQ ID NO 6732; 1701pp + Sequence Listing; English.
 XX
 CC The invention relates to novel genes (ABAI1004-ABAI21534) and proteins
 CC (ABAI4678-ABAI8001) useful for preventing, treating or ameliorating
 CC medical conditions e.g. by protein or gene therapy. The genes are
 CC isolated from a range of human tissues disclosed in the specification.
 CC The nucleic acids, proteins, antibodies and antagonists are useful in
 CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
 CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
 CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
 CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
 CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
 CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
 CC cardiovascular disorders such as myocardial ischaemias; (d) wound healing
 CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
 CC infectious diseases such as viral, bacterial, fungal and parasitic
 CC infections. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 798 BP; 227 A; 184 C; 174 G; 213 T; 0 U; 0 Other;

 Query Match 87.2%; Score 21.8; DB 5; Length 798;
 Best Local Similarity 92.0%; Pred. No. 4.5;
 Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

 QY 1 AAAAAGCTGCATCCCAATGTTAT 25
 DB 764 AAAAAGCTGCATCCCAATGTTAT 788

 RESULT 10
 ACA46717
 ID ACA46717 standard; DNA; 1914 BP.
 XX
 AC ACA46717;
 XX
 DT 19-JUN-2003 (first entry)
 XX
 DE Prokaryotic essential gene #28374.
 XX
 KW Antisense; ds; prokaryotic essential gene; cell proliferation;
 KW drug design; gene.
 XX
 OS Staphylococcus epidermidis.

XX WO200277183-A2.
 XX 03-OCT-2002.
 XX
 PF 21-MAR-2002; 2002WO-US009107.
 XX
 PR 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX
 XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX
 DR WPI; 2003-029926/02.
 DR P-PSDB; ABU42847.
 XX
 PT New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX
 PS Claim 14; SEQ ID NO 34587; 1766pp; English.
 XX
 CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing of the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
 CC prokaryotic essential genes. Note: The sequence data for this patent did
 CC not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 1914 BP; 648 A; 321 C; 323 G; 622 T; 0 U; 0 Other;

Query Match 81.6%; Score 20.4; DB 8; Length 1914;
 Best Local Similarity 95.5%; Pred. No. 23;
 Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

 QY 3 AAAACTGCATCCCAATGTTAT 24
 DB 1441 AAAACTGCATCCCAATGTTAT 1462

 RESULT 11
 AAH53037
 ID AAH53037 standard; DNA; 1917 BP.

XX AAH53037;
AC
XX
XX 03-SEP-2001 (first entry)
XX
DE S. epidermidis open reading frame nucleotide sequence SEQ ID NO:1467.
XX
XX Staphylococcus epidermidis SR1 strain; infection; diagnosis; vaccination;
KW endocarditis; ds.
XX
XX Staphylococcus epidermidis.
OS
XX
XX WO200134809-A2.
PN
XX
XX 17-MAY-2001.
PD
XX
XX 09-NOV-2000; 2000WO-US030782.
PF
XX
XX 09-NOV-1999; 99US-0164258P.
PR
XX (GLAX) GLAXO GROUP LTD.
PA
XX
XX Kimmerly WJ;
PI
XX
XX WPI; 2001-316495/33.
DR
XX P-PSDB; AAG82187.
DR
XX
XX Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
PT useful for vaccinating against infections, e.g. endocarditis.
PT
XX
XX Claim 8; Page 414-415; 2188pp; English.
PS
XX
XX AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
CC (II), given in AAG81454 to AAG83120, from Staphylococcus epidermidis. (I)
CC and (II) can have antibacterial activity and therefore can be used in
CC vaccination. The nucleic acids (I) may be used to produce the S.
CC epidermidis polypeptides (II) via the production of vectors containing
CC them which are used to produce hosts cells which express the
CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
CC used to vaccinate subjects and to raise antibodies against the bacteria.
CC The polypeptides may also be used to assay for other inhibitors of their
CC activity and therefore identify compounds that may be used for the
CC treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
CC AAH5090 represent specifically claimed S. epidermidis genomic DNA
CC polynucleotide sequences from the present invention. AAH55091 to AAH5098
CC represent oligonucleotide sequences and primers which are used in the
CC exemplification of the present invention. N.B. The present invention
CC specifically claims all the polynucleotide sequences given in the
CC sequence listing of the present specification, however the sequence
CC listing only goes up to SEQ ID NO:4454 so even though sequences are given
CC in the disclosure for SEQ ID NO:4465 to 4472, no sequences are present
CC for SEQ ID NO:4455 to 4464
XX
XX Sequence 1917 BP; 650 A; 321 C; 323 G; 623 T; 0 U; 0 Other;
SQ
Query Match 81.6%; Score 20.4; DB 4; Length 1917;
Best Local Similarity 95.5%; Pred. No. 23;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 3 AAAAAGTGCATCCCAATGTTAT 24
Db 1441 AAAAAGTGCATCCCAATGTTAT 1462
RESULT 12
ABN92144
ID ABN92144 standard; DNA; 1953 BP.
XX
AC ABN92144;
XX
XX 24-JUL-2002 (first entry)
DT
XX
XX Staphylococcus epidermidis ORF nucleic acid sequence SEQ ID NO:1607.

XX Staphylococcus epidermidis; open reading frame; ORF; bacterial infection;
KW antibacterial; gene therapy; gene; ds.
XX
XX Staphylococcus epidermidis.
OS
XX
XX US6380370-B1.
PN
XX
XX 30-APR-2002.
PD
XX
XX 13-AUG-1998; 98US-00134001.
PF
XX
XX 14-AUG-1997; 97US-0055779P.
PR
XX 08-NOV-1997; 97US-0064964P.
PR
XX (GENO-) GENOME THERAPEUTICS CORP.
PA
XX
XX Doucette-Stamm LA, Bush D;
PI
XX
XX WPI; 2002-381255/41.
DR
XX P-PSDB; ABP39599.
DR
XX
XX Novel isolated nucleic acid encoding a Staphylococcus epidermis
PT polypeptide, useful for diagnosing and treating bacterial infections.
PT
XX
XX Disclosure; SEQ ID NO 1607; 267pp; English.
PS
XX
XX ABN90538 to ABN93374 represent Staphylococcus epidermidis open reading
CC frame (ORF) nucleic acid sequences which encode the amino acid sequences
CC given in ABP35124 to ABP37960. The S. epidermidis sequences have
CC antibacterial activity and can be used in gene therapy. The sequences can
CC also be used in the diagnosis and treatment of bacterial infections,
CC particularly S. epidermidis infections. The sequences can be used to
CC screen for compounds able to interfere with the S. epidermidis life cycle
CC or inhibit S. epidermidis infection. N.B. The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from the USPTO web site
XX
XX Sequence 1953 BP; 663 A; 320 C; 338 G; 632 T; 0 U; 0 Other;
SQ
Query Match 81.6%; Score 20.4; DB 6; Length 1953;
Best Local Similarity 95.5%; Pred. No. 23;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 3 AAAAAGTGCATCCCAATGTTAT 24
Db 1477 AAAAAGTGCATCCCAATGTTAT 1498
RESULT 13
ADS01836
ID ADS01836 standard; DNA; 1953 BP.
XX
XX ADS01836;
AC
XX
XX 04-NOV-2004 (first entry)
DT
XX
XX Staphylococcus epidermis polynucleotide seqid 1131.
DE
XX
XX antibacterial; vaccine; antisense therapy; Staphylococcus epidermidis;
KW recombinant expression vector; infection; computer readable medium;
KW computer based system; gene; ds.
XX
XX Staphylococcus epidermidis.
OS
XX
XX US2004147734-A1.
PN
XX
XX 29-JUL-2004.
PD
XX
XX 01-DEC-2003; 2003US-00724972.
PF
XX
XX 08-NOV-1997; 97US-0064964P.
PR
XX 13-AUG-1998; 98US-00134001.
PR

FT	variation	5183		FT	
FT		/tag= d	"Single nucleotide polymorphism (SNP)"	FT	/tag= ab
FT		/standard_name=		FT	/standard_name= "Single nucleotide polymorphism (SNP)"
FT	variation	6839		FT	35131
FT		/tag= e	"Single nucleotide polymorphism (SNP)"	FT	/tag= ac
FT		/standard_name=		FT	/standard_name= "Single nucleotide polymorphism (SNP)"
FT	variation	7016		FT	35511
FT		/tag= f	"Single nucleotide polymorphism (SNP)"	FT	/tag= ad
FT		/standard_name=		FT	/standard_name= "Single nucleotide polymorphism (SNP)"
FT	variation	7237		FT	35708
FT		/tag= g	"Single nucleotide polymorphism (SNP)"	FT	/tag= ae
FT		/standard_name=		FT	/standard_name= "Single nucleotide polymorphism (SNP)"
FT	variation	8190		FT	36071
FT		/tag= h	"Single nucleotide polymorphism (SNP)"	FT	/tag= af
FT		/standard_name=		FT	/standard_name= "Single nucleotide polymorphism (SNP)"
FT	variation	8357		FT	36824
FT		/tag= i	"Single nucleotide polymorphism (SNP)"	FT	/tag= ag
FT		/standard_name=		FT	/standard_name= "Single nucleotide polymorphism (SNP)"
FT	variation	8577		FT	36959
FT		/tag= j	"Single nucleotide polymorphism (SNP)"	FT	/tag= ah
FT		/standard_name=		FT	/standard_name= "Single nucleotide polymorphism (SNP)"
FT	variation	15999		FT	37335
FT		/tag= k	"Single nucleotide polymorphism (SNP)"	FT	/tag= ai
FT		/standard_name=		FT	/standard_name= "Single nucleotide polymorphism (SNP)"
FT	variation	17722		FT	37343
FT		/tag= l	"Single nucleotide polymorphism (SNP)"	FT	/tag= aj
FT		/standard_name=		FT	/standard_name= "Single nucleotide polymorphism (SNP)"
FT	variation	18543		FT	38507
FT		/tag= m	"Single nucleotide polymorphism (SNP)"	FT	/tag= ak
FT		/standard_name=		FT	/standard_name= "Single nucleotide polymorphism (SNP)"
FT	variation	21639		FT	38554
FT		/tag= n	"Single nucleotide polymorphism (SNP)"	FT	/tag= al
FT		/standard_name=		FT	/standard_name= "Single nucleotide polymorphism (SNP)"
FT	variation	24753		FT	38624
FT		/tag= o	"Single nucleotide polymorphism (SNP)"	FT	/tag= am
FT		/standard_name=		FT	/standard_name= "Single nucleotide polymorphism (SNP)"
FT	variation	26289		FT	38690
FT		/tag= p	"Single nucleotide polymorphism (SNP)"	FT	/tag= an
FT		/standard_name=		FT	/standard_name= "Single nucleotide polymorphism (SNP)"
FT	variation	27120		FT	38744
FT		/tag= q	"Single nucleotide polymorphism (SNP)"	FT	/tag= ao
FT		/standard_name=		FT	/standard_name= "Single nucleotide polymorphism (SNP)"
FT	variation	28737		FT	38756
FT		/tag= r	"Single nucleotide polymorphism (SNP)"	FT	/tag= ap
FT		/standard_name=		FT	/standard_name= "Single nucleotide polymorphism (SNP)"
FT	variation	29562		FT	39525
FT		/tag= s	"Single nucleotide polymorphism (SNP)"	FT	/tag= aq
FT		/standard_name=		FT	/standard_name= "Single nucleotide polymorphism (SNP)"
FT	variation	30078		FT	40682
FT		/tag= t	"Single nucleotide polymorphism (SNP)"	FT	/tag= ar
FT		/standard_name=		FT	/standard_name= "Single nucleotide polymorphism (SNP)"
FT	variation	30974		FT	41181
FT		/tag= u	"Single nucleotide polymorphism (SNP)"	FT	/tag= as
FT		/standard_name=		FT	/standard_name= "Single nucleotide polymorphism (SNP)"
FT	variation	31545		FT	41414
FT		/tag= v	"Single nucleotide polymorphism (SNP)"	FT	/tag= at
FT		/standard_name=		FT	/standard_name= "Single nucleotide polymorphism (SNP)"
FT	variation	31779		FT	41657
FT		/tag= w	"Single nucleotide polymorphism (SNP)"	FT	/tag= au
FT		/standard_name=		FT	/standard_name= "Single nucleotide polymorphism (SNP)"
FT	variation	34062		FT	42821
FT		/tag= x	"Single nucleotide polymorphism (SNP)"	FT	/tag= av
FT		/standard_name=		FT	/standard_name= "Single nucleotide polymorphism (SNP)"
FT	variation	34831		FT	45475
FT		/tag= y	"Single nucleotide polymorphism (SNP)"	FT	/tag= aw
FT		/standard_name=		FT	/standard_name= "Single nucleotide polymorphism (SNP)"
FT	variation	34906		FT	45611
FT		/tag= z	"Single nucleotide polymorphism (SNP)"	FT	/tag= ax
FT		/standard_name=		FT	/standard_name= "Single nucleotide polymorphism (SNP)"
FT	variation	34934		FT	45763
FT		/tag= aa	"Single nucleotide polymorphism (SNP)"	FT	/tag= ay
FT		/standard_name=		FT	/standard_name= "Single nucleotide polymorphism (SNP)"
FT	variation	34938		FT	46038
FT		/tag= az		FT	/tag= az

CC markers for spermatogonial cells, for identifying genes or proteins
CC characteristic of male infertility, diagnosing or aiding in the diagnosis
CC of infertility in men, and for contraception in which sperm production or
CC sperm count is reduced or defective sperm is produced. Antibodies to
CC reproductive-specific proteins are useful for determining the presence of
CC these proteins in a sample obtained from a man being assessed for
CC infertility, for identifying the expression of genes in particular cell
CC type or particular developmental stage, for studies of spermatogenesis,
CC and for immunofluorescence of germ cells or in Western blots for
CC assessing the presence of the protein the antibody binds. The sequences
CC of the invention are also useful for treating disorders of reduced sperm
CC count, and for increasing sperm count and/or sperm activity. The nucleic
CC acids of the invention are useful in gene therapy. AAS13648-AAS13671
CC represent nucleic acid sequences encoding for the human reproduction-
CC specific proteins of the present invention

XX SQ Sequence 53552 BP; 14743 A; 13420 C; 12475 G; 12914 T; 0 U; 0 Other;
Query Match 75.2%; Score 18.8; DB 5; Length 53552;
Best Local Similarity 90.9%; Pred. No. 2.1e+02;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 ABACTGCATCCCAATGTTATG 25
||||| ||||||| ||||| |||||
Db 29520 AACTTCATCCCAATGTTATG 29541

RESULT 18
ABT11173
ID ABT11173 standard; DNA; 168174 BP.
XX AC ABT11173;
XX DT 05-DEC-2002 (first entry)
XX DE Human 5-lipoxygenase gene related DNA sequence SEQ ID No 63.
XX KW Human; polymorphic region; 5-lipoxygenase; 5-LO gene; asthma; bronchitis;
KW sinusitis; ulcerative colitis; nephritis; amyloidosis; sarcoidosis;
KW rheumatoid arthritis; scleroderma; lupus; non-allergic rhinitis;
KW polymyositis; Reiter's syndrome; psoriasis; pelvic inflammatory disease;
KW atopic; contact dermatitis; forensic medicine; paternity testing; enzyme;
ds.
XX OS Homo sapiens.
XX WO200262825-A2.
XX PD 15-AUG-2002.
XX PF 07-FEB-2002; 2002WO-US003546.
XX PR 08-FEB-2001; 2001US-0267515P.
XX PR 21-AUG-2001; 2001US-0314248P.
XX PA (MILL-) MILLENNIUM PHARM INC.
XX PI Barnes G, Meyer J;
XX DR WPI; 2002-627522/67.
XX PT New isolated nucleic acid molecule with an allelic variant of a
XX polymorphic region of an 5-LO gene, useful for diagnosing and/or
XX prognosticating disorders associated with an aberrant inflammatory
XX response such as asthma.
XX PS Disclosure; Fig 4; 290pp; English.
XX CC The invention relates to an isolated human nucleic acid molecule
XX comprising an allelic variant of a polymorphic region of a 5-lipoxygenase
XX (5-LO) gene, where the allelic variant comprises one or more nucleotide
XX selected from any of 3, 20 or 21 base pair sequences, given in the
XX specification, or their complement. The compositions and methods of the
XX invention relates to an isolated human nucleic acid molecule
XX comprising an allelic variant of a polymorphic region of a 5-lipoxygenase
XX (5-LO) gene, where the allelic variant comprises one or more nucleotide
XX selected from any of 3, 20 or 21 base pair sequences, given in the
XX specification, or their complement. The compositions and methods of the

CC present invention are useful for diagnosing and/or prognosing disorders
CC associated with an aberrant inflammatory response such as asthma,
CC bronchitis, sinusitis, ulcerative colitis, nephritis, amyloidosis,
CC rheumatoid arthritis, sarcoidosis, scleroderma, lupus, non-allergic
CC rhinitis, polymyositis, Reiter's syndrome, psoriasis, pelvic inflammatory
CC disease, atopic and contact dermatitis. The nucleic acid molecules can
CC also be useful for identifying an individual amongst other individuals
CC from the same species for use in forensic medicine and paternity testing.
CC This polymucleotide sequence represents DNA relating to the human 5-
CC lipoxygenase (5-LO) gene of the invention

XX SQ Sequence 168174 BP; 46808 A; 36442 C; 36942 G; 46474 T; 0 U; 1508 Other;
Query Match 74.4%; Score 18.6; DB 6; Length 168174;
Best Local Similarity 84.0%; Pred. No. 3.1e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAGCTGCATCCCAATGTTATG 25
||||| ||||||| ||||| |||||
Db 91147 AAAGAACTGCATCCCTTAATGTTCTG 91171

RESULT 19
ABT11114
ID ABT11114 standard; DNA; 168273 BP.
XX AC ABT11114;
XX DT 05-DEC-2002 (first entry)
XX DE Human 5-lipoxygenase gene related DNA sequence SEQ ID No 2.
XX KW Human; polymorphic region; 5-lipoxygenase; 5-LO gene; asthma; bronchitis;
KW sinusitis; ulcerative colitis; nephritis; amyloidosis; sarcoidosis;
KW rheumatoid arthritis; scleroderma; lupus; non-allergic rhinitis;
KW polymyositis; Reiter's syndrome; psoriasis; pelvic inflammatory disease;
KW atopic; contact dermatitis; forensic medicine; paternity testing; enzyme;
ds.
XX OS Homo sapiens.
XX WO200262825-A2.
XX PD 15-AUG-2002.
XX PF 07-FEB-2002; 2002WO-US003546.
XX PR 08-FEB-2001; 2001US-0267515P.
XX PR 21-AUG-2001; 2001US-0314248P.
XX PA (MILL-) MILLENNIUM PHARM INC.
XX PI Barnes G, Meyer J;
XX DR WPI; 2002-627522/67.
XX PT New isolated nucleic acid molecule with an allelic variant of a
XX polymorphic region of an 5-LO gene, useful for diagnosing and/or
XX prognosticating disorders associated with an aberrant inflammatory
XX response such as asthma.
XX PS Disclosure; Fig 2; 290pp; English.
XX CC The invention relates to an isolated human nucleic acid molecule
XX comprising an allelic variant of a polymorphic region of a 5-lipoxygenase
XX (5-LO) gene, where the allelic variant comprises one or more nucleotide
XX selected from any of 3, 20 or 21 base pair sequences, given in the
XX specification, or their complement. The compositions and methods of the
XX present invention are useful for diagnosing and/or prognosing disorders
XX associated with an aberrant inflammatory response such as asthma,
XX bronchitis, sinusitis, ulcerative colitis, nephritis, amyloidosis,
XX rheumatoid arthritis, sarcoidosis, scleroderma, lupus, non-allergic
XX rhinitis, polymyositis, Reiter's syndrome, psoriasis, pelvic inflammatory

AD57983/c
ID ADE57983 standard; DNA; 521 BP.
XX
AC ADE57983;
XX
DT 29-JAN-2004 (first entry)
XX
DE Human gene AV701053, SEQ ID NO 3851.
XX
XX Human; ds; gene; pain; neuronal tissue; gene therapy;
KW spinal segmental nerve injury; chronic constriction injury; CCI;
KW spared nerve injury; SNI; Chung.
XX
XX Homo sapiens.
XX
XX WO2003016475-A2.
PN
XX 27-FEB-2003.
PD
XX 14-AUG-2002; 2002WO-US025765.
PF
XX 14-AUG-2001; 2001US-0312147P.
PR
XX 01-NOV-2001; 2001US-0346382P.
PR
XX 26-NOV-2001; 2001US-0333347P.
XX
XX (GEHO) GEN HOSPITAL CORP.
PA (FARB) BAYER AG.
XX
XX Woolf C, D'urso D, Befort K, Costigan M;
PI
XX WPI; 2003-268312/26.
DR GENBANK; AV701053.
XX
XX New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
PT
XX Claim 1; Page; 1017pp; English.
XX
XX The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC injury (e.g. spinal segmental nerve injury (SNI)) in an animal (e.g. gene
CC injury (CCI) and spared nerve injury (SNI)) in a human DNA (shown in Table 2 of the
CC specification) which encodes one of the polypeptides of the invention
CC which is differentially expressed during pain. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 521 BP; 180 A; 75 C; 122 G; 144 T; 0 U; 0 Other;
SQ
Query Match 72.8%; Score 18.2; DB 10; Length 521;
Best Local Similarity 87.0%; Pred. No. 2.1e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAAAGTGCATCCCAATGTTA 23
Db 440 AAAAAAAGTGCATCCCAATGTTA 418
RESULT 23
ADE60446/c
ID ADE60446 standard; DNA; 521 BP.
XX
AC ADE60446;
XX
DT 29-JAN-2004 (first entry)
XX
DE Human gene AV701053, SEQ ID NO 6355.
XX
XX Human; ds; gene; pain; neuronal tissue; gene therapy;
KW spinal segmental nerve injury; chronic constriction injury; CCI;
KW spared nerve injury; SNI; Chung.
XX
XX Homo sapiens.
XX
XX WO2003016475-A2.
PN
XX 27-FEB-2003.
PD
XX 14-AUG-2002; 2002WO-US025765.
PF
XX 14-AUG-2001; 2001US-0312147P.
PR
XX 01-NOV-2001; 2001US-0346382P.
PR
XX 26-NOV-2001; 2001US-0333347P.
XX
XX (GEHO) GEN HOSPITAL CORP.
PA (FARB) BAYER AG.
XX
XX Woolf C, D'urso D, Befort K, Costigan M;
PI
XX WPI; 2003-268312/26.
DR GENBANK; AV701053.
XX
XX New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
PT
XX Claim 1; Page; 1017pp; English.
XX
XX The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC injury (e.g. spinal segmental nerve injury (SNI)) in an animal (e.g. gene
CC injury (CCI) and spared nerve injury (SNI)) in a human DNA (shown in Table 2 of the
CC specification) which encodes one of the polypeptides of the invention
CC which is differentially expressed during pain. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
XX


```
SQ Sequence 521 BP; 180 A; 75 C; 122 G; 144 T; 0 U; 0 Other;
Query Match 72.8%; Score 18.2; DB 10; Length 521;
Best Local Similarity 87.0%; Pred. No. 2.1e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAAAGTCATCCCCCAATGTTA 23
   ||||| ||||| ||||| ||||| |||||
Db 440 AAAAAAAGTCATCCCCCAATATAA 418

RESULT 24
ADE60727/c
ID ADE60727 standard; DNA; 521 BP.
XX
AC ADE60727;
XX
DT 29-JAN-2004 (first entry)
XX
DE Human gene AV701053, SEQ ID NO 6639.
XX
KW Human; ds; gene; pain; neuronal tissue; gene therapy;
KW spinal segmental nerve injury; chronic constriction injury; CCI;
KW spared nerve injury; SNI; Chung.
XX
OS Homo sapiens.
XX
PN WO2003016475-A2.
XX
PD 27-FEB-2003.
XX
PF 14-AUG-2002; 2002WO-US025765.
XX
PR 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333347P.
XX
PA (GEHO ) GEN HOSPITAL CORP.
PA (FARB ) BAYER AG.
XX
PI Woolf C, D'urso D, Befort K, Costigan M;
XX
WPI; 2003-268312/26.
DR GENBANK; AV701053.
XX
PT New composition comprising two or more isolated polypeptides, useful for
preparing a medicament for treating pain in an animal.
XX
PS Claim 1; Page; 1017pp; English.
XX
CC The invention discloses a composition comprising two or more isolated rat
or human polynucleotides or a polynucleotide which represents a fragment,
derivative or allelic variation of the nucleic acid sequence. Also
claimed are a vector comprising the novel polynucleotide, a host cell
comprising the vector, a method for identifying a nucleotide sequence
which is differentially regulated in an animal subjected to pain and a
kit to perform the method, an array, a method for identifying an agent
that increases or decreases the expression of the polynucleotide sequence
that is differentially expressed in neuronal tissue of a first animal
subjected to pain, a method for identifying a compound which regulates
the expression of a polynucleotide sequence which is differentially
expressed in an animal subjected to pain, a method for identifying a
compound that regulates the activity of one or more of the
polynucleotides, a method for producing a pharmaceutical composition, a
method for identifying a compound or small molecule that regulates the
activity in an animal of one or more of the polypeptides given in the
specification, a method for identifying a compound useful in treating
pain and a pharmaceutical composition comprising the one or more
polypeptides or their antibodies. The polynucleotide or the compound that
modulates its activity is useful for preparing a medicament for treating
pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
therapy). The sequence presented is a human DNA (shown in Table 2 of the
specification) which encodes one of the polypeptides of the invention
which is differentially expressed during pain. Note: The sequence data
for this patent did not form part of the printed specification, but was
obtained in electronic form directly from WIPO at
ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 521 BP; 180 A; 75 C; 122 G; 144 T; 0 U; 0 Other;
Query Match 72.8%; Score 18.2; DB 10; Length 521;
Best Local Similarity 87.0%; Pred. No. 2.1e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAAAGTCATCCCCCAATGTTA 23
   ||||| ||||| ||||| ||||| |||||
Db 440 AAAAAAAGTCATCCCCCAATATAA 418

RESULT 25
ADE60731/c
ID ADE60731 standard; DNA; 521 BP.
XX
AC ADE60731;
XX
DT 29-JAN-2004 (first entry)
XX
DE Human gene AV701053, SEQ ID NO 6643.
XX
KW Human; ds; gene; pain; neuronal tissue; gene therapy;
KW spinal segmental nerve injury; chronic constriction injury; CCI;
KW spared nerve injury; SNI; Chung.
XX
OS Homo sapiens.
XX
PN WO2003016475-A2.
XX
PD 27-FEB-2003.
XX
PF 14-AUG-2002; 2002WO-US025765.
XX
PR 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333347P.
XX
PA (GEHO ) GEN HOSPITAL CORP.
PA (FARB ) BAYER AG.
XX
PI Woolf C, D'urso D, Befort K, Costigan M;
XX
WPI; 2003-268312/26.
DR GENBANK; AV701053.
XX
PT New composition comprising two or more isolated polypeptides, useful for
preparing a medicament for treating pain in an animal.
XX
PS Claim 1; Page; 1017pp; English.
XX
CC The invention discloses a composition comprising two or more isolated rat
or human polynucleotides or a polynucleotide which represents a fragment,
derivative or allelic variation of the nucleic acid sequence. Also
claimed are a vector comprising the novel polynucleotide, a host cell
comprising the vector, a method for identifying a nucleotide sequence
which is differentially regulated in an animal subjected to pain and a
kit to perform the method, an array, a method for identifying an agent
that increases or decreases the expression of the polynucleotide sequence
that is differentially expressed in neuronal tissue of a first animal
subjected to pain, a method for identifying a compound which regulates
the expression of a polynucleotide sequence which is differentially
expressed in an animal subjected to pain, a method for identifying a
compound that regulates the activity of one or more of the
polynucleotides, a method for producing a pharmaceutical composition, a
method for identifying a compound or small molecule that regulates the
activity in an animal of one or more of the polypeptides given in the
specification, a method for identifying a compound useful in treating
pain and a pharmaceutical composition comprising the one or more
polypeptides or their antibodies. The polynucleotide or the compound that
modulates its activity is useful for preparing a medicament for treating
pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
therapy). The sequence presented is a human DNA (shown in Table 2 of the
```

CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a human DNA (shown in Table 2 of the
CC specification) which encodes one of the polypeptides of the invention
CC which is differentially expressed during pain. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 521 BP; 180 A; 75 C; 122 G; 144 T; 0 U; 0 Other;

Query Match 72.8%; Score 18.2; DB 10; Length 521;
Best Local Similarity 87.0%; Pred. No. 2.1e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAAAGTCATCCCAATGTTA 23
|||||||
Dd 440 AAAAAAAGTCATCCCAATATAA 418

RESULT 26

ACN43954
ID ACN43954 standard; DNA; 95683 BP.

XX ACN43954;

DT 18-NOV-2004 (first entry)

XX Human genomic sequence hCG37570.

XX Cytostatic; carcinoma; lymphoma; cancer; human; gene; ss.

XX Homo sapiens.

XX WO2003073826-A2.

XX 12-SEP-2003.

XX 28-FEB-2003; 2003WO-US006235.

XX 01-MAR-2002; 2002US-00087192.

XX (SAGR-) SAGRES DISCOVERY.

XX Morris DW;

XX WPI; 2003-328604/31.

XX Recombinant nucleic acid useful for diagnosis and treatment of carcinoma
PT comprises a nucleotide sequence.

XX Claim 1; SEQ ID NO 160; Opp; English.

XX The present invention relates to novel DNA and protein sequences which
CC are associated with carcinomas. The sequences are useful for: (i) for
CC screening drug candidates; (ii) for screening of bioactive agent capable
CC of binding to Carcino Associated Protein (CAP); (iii) for screening of
CC a bioactive agent capable of modulating the activity of CAP; (iv) for
CC evaluating the effect of a candidate carcinoma drug; (v) for diagnosing
CC carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating
CC carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip;
CC (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for
CC determining Carcino Associated (CA) gene copy number. In addition, the
CC CA genes are useful as DNA vaccines and the CAP are useful as markers of
CC carcinoma including lymphoma. The present sequence is one such CA coding
CC sequence. Note: This patent is an equivalent to basic patent
CC US2002182586A1, for which no sequence data was published

XX Sequence 95683 BP; 28002 A; 20734 C; 20466 G; 26481 T; 0 U; 0 Other;

Query Match 72.8%; Score 18.2; DB 11; Length 95683;
Best Local Similarity 87.0%; Pred. No. 4.4e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAAAGTCATCCCAATGTTAT 24
|||||||
Dd 8156 AAAAAAGTCATCCCAATGTTAT 8178

RESULT 27

ADL13855
ID ADL13855 standard; DNA; 152048 BP.

XX ADL13855;

XX 06-MAY-2004 (first entry)

XX Osteoarthritis-associated polymorphic nucleotide #387.

XX ds; gene; osteopathic; antiinflammatory; antiarthritic; gene therapy;
KW joint space narrowing; osteophyte development; joint pain;
KW osteoarthritis; SNP; single nucleotide polymorphism.

XX Homo sapiens.

XX WO2003054166-A2.

XX 03-JUL-2003.

XX 19-DEC-2002; 2002WO-US041225.

XX 20-DEC-2001; 2001US-0342603P.

XX (INCY-) INCYTE GENOMICS INC.

XX Jones KA, Schafer A;

XX WPI; 2003-559141/52.

XX Determining susceptibility of an individual to joint space narrowing,
PT osteophyte development and/or joint pain comprises identifying whether
PT the individual has at least one polymorphism in a polynucleotide encoding
PT a protein.

XX Disclosure; SEQ ID NO 387; 297pp; English.

XX The invention relates to a method of determining susceptibility of an
CC individual to joint space narrowing and/or osteophyte development and/or
CC joint pain comprising identifying whether the individual has at least one
CC polymorphism in a polynucleotide encoding at least one of the protein
CC listed in the specification. The methods, composition and agent are
CC useful for modulating the susceptibility of an individual to joint space
CC narrowing and/or osteophyte development and/or joint pain that is
CC associated with a disease, preferably osteoarthritis. The cell line and
CC the non-human animal are useful for screening for an agent for diagnosing
CC an individual having susceptibility to joint space narrowing and/or
CC osteophyte development and/or joint pain. This sequence corresponds to
CC the polynucleotide encoding a protein listed in the specification. (Note:
CC The sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences).

XX Sequence 152048 BP; 45197 A; 32444 C; 31900 G; 42505 T; 0 U; 2 Other;

Query Match 72.8%; Score 18.2; DB 10; Length 152048;

Best Local Similarity 87.0%; Pred. No. 4.7e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAAGTCATCCCAATGTTA 23

Dd 95029 AAAAAAGTCATCCCAATGTTA 95051

```
RESULT 28
ABQ31619
ID ABQ31619 standard; DNA; 578 BP.
XX
AC ABQ31619;
XX
DT 12-JUL-2002 (first entry)
XX
DE Oligonucleotide for detecting cytosine methylation SEQ ID NO 18210.
XX
KW Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
KW drug; side effect; cancer; central nervous system; cardiovascular;
KW gastrointestinal; respiratory system; single nucleotide polymorphism;
KW SNP; cell differentiation; ds.
XX
OS Homo sapiens.
XX
PN WO200218632-A2.
XX
PD 07-MAR-2002.
XX
PF 01-SEP-2001; 2001WO-EP010074.
XX
PR 01-SEP-2000; 2000DE-01043826.
XX
PP 05-SEP-2000; 2000DE-01044543.
XX
PA (EPiG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K, Guetig D;
XX
PP WI; 2002-371829/40.
XX
PT Determining the degree of cytosine methylation in genomic DNA, useful for
PT diagnosis and prognosis, comprises selective hybridization of amplicons
PT from chemically treated DNA.
XX
PS Claim 12; 56pp + Sequence Listing; 56pp; German.
XX
CC This invention describes a novel method for determining the degree of
CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a
CC genomic sample of DNA. The sample is treated chemically to convert
CC cytosine (C) but not methylated C, to uracil, then part of the genomic
CC DNA that contains the target C is amplified to form a labeled amplicon.
CC The amplicon is hybridised to two classes, each with at least one member,
CC of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the
CC degree of hybridisation to both classes is determined from the label on
CC the amplicon. From the ratio of labels hybridised to the two classes of
CC oligomers, the degree of methylation is calculated. The method is used:
CC (i) for diagnosis and/or prognosis of side effects of therapeutic drugs
CC and of a wide range of diseases, e.g. cancer, disorders of the central
CC nervous, cardiovascular, gastrointestinal and respiratory systems etc.,
CC particularly by detecting mutations or single nucleotide polymorphisms
CC (SNP's); and (ii) for differentiation of cell or tissue types and for
CC investigating cell differentiation. The method allows the methylation
CC status of many C residues to be determined simultaneously. ABQ13410-
CC ABQ54121 represent genomic DNA sequences used to illustrate the method
CC for determining the degree of cytosine methylation described in the
CC disclosure of the invention
XX
SQ Sequence 578 BP; 213 A; 230 C; 77 G; 58 T; 0 U; 0 Other;
XX
Query Match 71.2%; Score 17.8; DB 6; Length 578;
Best Local Similarity 90.5%; Pred. No. 3.3e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 AAAAACTGCATCCCAATGT 21
Db 176 AAAAAACGCATCCCAATTT 196
RESULT 29
ABQ31618/c
ID -ABQ31618 standard; DNA; 578 BP.
XX
```

```
XX
AC ABQ31618;
XX
DT 12-JUL-2002 (first entry)
XX
DE Oligonucleotide for detecting cytosine methylation SEQ ID NO 18209.
XX
KW Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
KW drug; side effect; cancer; central nervous system; cardiovascular;
KW gastrointestinal; respiratory system; single nucleotide polymorphism;
KW SNP; cell differentiation; ds.
XX
OS Homo sapiens.
XX
PN WO200218632-A2.
XX
PD 07-MAR-2002.
XX
PF 01-SEP-2001; 2001WO-EP010074.
XX
PR 01-SEP-2000; 2000DE-01043826.
XX
PP 05-SEP-2000; 2000DE-01044543.
XX
PA (EPiG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K, Guetig D;
XX
PP WI; 2002-371829/40.
XX
PT Determining the degree of cytosine methylation in genomic DNA, useful for
PT diagnosis and prognosis, comprises selective hybridization of amplicons
PT from chemically treated DNA.
XX
PS Claim 12; 56pp + Sequence Listing; 56pp; German.
XX
CC This invention describes a novel method for determining the degree of
CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a
CC genomic sample of DNA. The sample is treated chemically to convert
CC cytosine (C) but not methylated C, to uracil, then part of the genomic
CC DNA that contains the target C is amplified to form a labeled amplicon.
CC The amplicon is hybridised to two classes, each with at least one member,
CC of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the
CC degree of hybridisation to both classes is determined from the label on
CC the amplicon. From the ratio of labels hybridised to the two classes of
CC oligomers, the degree of methylation is calculated. The method is used:
CC (i) for diagnosis and/or prognosis of side effects of therapeutic drugs
CC and of a wide range of diseases, e.g. cancer, disorders of the central
CC nervous, cardiovascular, gastrointestinal and respiratory systems etc.,
CC particularly by detecting mutations or single nucleotide polymorphisms
CC (SNP's); and (ii) for differentiation of cell or tissue types and for
CC investigating cell differentiation. The method allows the methylation
CC status of many C residues to be determined simultaneously. ABQ13410-
CC ABQ54121 represent genomic DNA sequences used to illustrate the method
CC for determining the degree of cytosine methylation described in the
CC disclosure of the invention
XX
SQ Sequence 578 BP; 58 A; 77 C; 230 G; 213 T; 0 U; 0 Other;
XX
Query Match 71.2%; Score 17.8; DB 6; Length 578;
Best Local Similarity 90.5%; Pred. No. 3.3e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 AAAAACTGCATCCCAATGT 21
Db 403 AAAAAACGCATCCCAATTT 383
RESULT 30
ACA24290/c
ID ACA24290 standard; DNA; 1083 BP.
XX
AC ACA24290;
XX
```

DT 19-JUN-2003 (first entry)

XX DE Prokaryotic essential gene #5947.

XX KW Antisense; ds; prokaryotic essential gene; cell proliferation;

XX KW drug design; gene.

XX OS Bacteroides fragilis.

XX PN WO200277183-A2.

XX XX 03-OCT-2002.

XX XX 21-MAR-2002; 2002WO-US009107.

XX XX 21-MAR-2001; 2001US-00815242.

XX PR 06-SEP-2001; 2001US-00948993.

XX PR 25-OCT-2001; 2001US-0342923P.

XX PR 08-FEB-2002; 2002US-00072851.

XX PR 06-MAR-2002; 2002US-0362699P.

XX XX (ELIT-) ELITRA PHARM INC.

XX XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;

XX PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX XX WPI; 2003-029926/02.

XX DR P-PSDB; ABU20420.

XX XX New antisense nucleic acids, useful for identifying proteins or screening

XX PT for homologous nucleic acids required for cellular proliferation to

XX PT isolate candidate molecules for rational drug discovery programs.

XX XX Claim 14; SEQ ID NO 12160; 1766pp; English.

XX XX The invention relates to an isolated nucleic acid comprising any one of

XX CC the 6213 antisense sequences given in the specification where expression

XX CC of the nucleic acid inhibits proliferation of a cell. Also included are:

XX CC (1) a vector comprising a promoter operably linked to the nucleic acid

XX CC encoding a polypeptide whose expression is inhibited by the antisense

XX CC nucleic acid; (2) a host cell containing the vector; (3) an isolated

XX CC polypeptide or its fragment whose expression is inhibited by the

XX CC antisense nucleic acid; (4) an antibody capable of specifically binding

XX CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular

XX CC proliferation or the activity of a gene in an operon required for

XX CC the gene product or that has an activity against a biological pathway

XX CC required for proliferation, or that inhibits cellular proliferation; (8)

XX CC identifying a gene required for cellular proliferation or the biological

XX CC pathway in which a proliferation-required gene or its gene product lies

XX CC or a gene on which the test compound that inhibits proliferation of an

XX CC organism acts; (9) manufacturing an antibiotic; (10) profiling a

XX CC compound's activity; (11) a culture comprising strains in which the gene

XX CC product is overexpressed or underexpressed; (12) determining the extent

XX CC to which each of the strains is present in a culture or collection of

XX CC strains; or (13) identifying the target of a compound that inhibits the

XX CC proliferation of an organism. The antisense nucleic acids are useful for

XX CC identifying proteins or screening for homologous nucleic acids required

XX CC for cellular proliferation to isolate candidate molecules for rational

XX CC drug discovery programs, or for screening homologous nucleic acids

XX CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,

XX CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target

XX CC prokaryotic essential genes. Note: The sequence data for this patent did

XX CC not form part of the printed specification, but was obtained in

XX CC electronic format directly from WIPO at

XX CC ftp.wipo.int/pub/published_pct_sequences

XX XX Sequence 1083 BP; 311 A; 180 C; 274 G; 318 T; 0 U; 0 Other;

SQ

Query Match 71.2%; Score 17.8; DB 8; Length 1083;

Best Local Similarity 90.5%; Pred. No. 3.6e+02;

Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AAAAAGTCATCCCCCAATGTT 22

DB 633 AAAAAGTCATCCCCCAATCTT 613

RESULT 31

Continuation (34 of 35) of ABB39175 from base 3300001 (L. pneumophila DNA SEQ ID NO 3507

WP Sequence split into 35 fragments LOCUS ABB39175 Accession Aeb39175

WP Fragment Name Begin End

WP ABB39175_00 1 110000

WP ABB39175_01 100001 210000

WP ABB39175_02 200001 310000

WP ABB39175_03 300001 410000

WP ABB39175_04 400001 510000

WP ABB39175_05 500001 610000

WP ABB39175_06 600001 710000

WP ABB39175_07 700001 810000

WP ABB39175_08 800001 910000

WP ABB39175_09 900001 1010000

WP ABB39175_10 1000001 1110000

WP ABB39175_11 1100001 1210000

WP ABB39175_12 1200001 1310000

WP ABB39175_13 1300001 1410000

WP ABB39175_14 1400001 1510000

WP ABB39175_15 1500001 1610000

WP ABB39175_16 1600001 1710000

WP ABB39175_17 1700001 1810000

WP ABB39175_18 1800001 1910000

WP ABB39175_19 1900001 2010000

WP ABB39175_20 2000001 2110000

WP ABB39175_21 2100001 2210000

WP ABB39175_22 2200001 2310000

WP ABB39175_23 2300001 2410000

WP ABB39175_24 2400001 2510000

WP ABB39175_25 2500001 2610000

WP ABB39175_26 2600001 2710000

WP ABB39175_27 2700001 2810000

WP ABB39175_28 2800001 2910000

WP ABB39175_29 2900001 3010000

WP ABB39175_30 3000001 3110000

WP ABB39175_31 3100001 3210000

WP ABB39175_32 3200001 3310000

WP ABB39175_33 3300001 3410000

WP ABB39175_34 3400001 3503610

Query Match 71.2%; Score 17.8; DB 14; Length 110000;

Best Local Similarity 90.5%; Pred. No. 6.9e+02;

Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 AACTGCATCCCCCAATGTTATG 25

DB 12430 AACTGAATCCCCCAATGCTATG 12410

RESULT 32

AEBS35716

ID AEBS35716 standard; DNA; 162147 BP.

XX AEBS35716;

XX 08-SEP-2005 (first entry)

XX L. pneumophila DNA SEQ ID NO 48.

XX detection; infection; Antibacterial; Vaccine; ds; gene.

XX Legionella pneumophila.

XX WO2005049642-A2.

XX 02-JUN-2005.

```
PF 23-SEP-2004; 2004WO-IB003578.
XX
PR 21-NOV-2003; 2003FR-00013687.
XX
XX (INSP ) INST PASTEUR.
PA (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.
PA (UPLY-) UNIV LYON 1 BERNARD CLAUDE.
PA (CNRS ) CNRS CENT NAT RECH SCI.
XX
XX Buchrieser C, Tichit M, Etienne J, Ma L, Cazalet C, Glaser P;
PI Rusniok C, Bouchier C, Zidane N, Magnier A, Kunst F, Vandenesch F;
PI Jarraud S;
XX
XX WPI; 2005-388305/40.
XX
XX New genome of Legionella pneumophila Paris strain and derived
PT polypeptides, useful for detection or identification of the strain and
PT for treatment and prevention of infections.
XX
XX Claim 1; SEQ ID NO 48; 660pp; English.
XX
XX The invention relates to an isolated or purified nucleotide sequences (I)
CC from Legionella pneumophila Paris strain. (II), and their related
CC sequences or fragments, are useful as primers and probes for detection
CC and amplification, including differentiation between the Paris and
CC Philadelphia strains of Legionella pneumophila and to prepare recombinant
CC (hybrid) polypeptides (II). (II) are also useful for preparation of
CC specific antibodies (Ab), also used for detection/identification of
CC Legionella, and some (I), specifically those involved in synthesis of
CC surface proteins, are targets for identification of inhibitors. (II), or
CC vectors that contain (I), are useful as vaccines and immunogenic
CC compositions, for treatment and prevention of infections by L.
CC pneumophila. The present sequence represents a L. pneumophila DNA.
XX
XX Sequence 162147 BP; 48444 A; 30180 C; 33920 G; 49593 T; 0 U; 10 Other;
SQ
Query Match 71.2%; Score 17.8; DB 14; Length 162147;
Best Local Similarity 90.5%; Pred. No. 7.2e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 5 AACTGCATCCCAATGTTATG 25
Db 141886 AACTGAATCCCAATGCTATG 141906
|||||
RESULT 33
ACH39621/c
ID ACH39621 standard; cdna; 461 BP.
XX
XX ACH39621;
XX
XX 13-OCT-2003 (first entry)
XX
XX Human foetal brain cdna #988.
XX
XX Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
KW genome mapping; biodiversity; genetic disorder.
XX
XX Homo sapiens.
XX
XX US2003073623-A1.
XX
XX 17-APR-2003.
XX
XX 30-JUL-2001; 2001US-00918995.
XX
XX 30-JUL-2001; 2001US-00918995.
XX
XX (DRMA/) DRMANAC R T.
PA (LABA/) LABAT I.
PA (STAC/) STACHE-CRAIN B.
PA (DICK/) DICKSON M C.
PA (JONE/) JONES L W.
XX
```

```
XX Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;
XX WPI; 2003-615964/58.
XX
XX New polynucleotide sequences obtained from various cDNA libraries, useful
PT as hybridization probes, as oligomers for PCR, for chromosome and gene
PT mapping, in the recombinant production of protein, or in generating
PT antisense DNA or RNA.
XX
XX Claim 1; SEQ ID NO 26833; 44pp; English.
XX
XX The invention relates to an isolated polynucleotide comprising any one of
CC 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
CC determined by the technique of SBH (sequencing by hybridisation). Also
CC included is a purified polypeptide comprising a sequence corresponding to
CC a reading frame of the novel polynucleotide. The nucleic acid sequences
CC are useful in diagnostics as expressed sequence tags (EST) for
CC identifying expressed genes or for physical mapping of the human genome,
CC in forensics, in assessing biodiversity, or in identifying mutations
CC responsible for genetic disorders and other traits. The nucleotide
CC sequences are also useful as hybridisation probes, as oligomers for PCR,
CC for chromosome and gene mapping, in the recombinant production of
CC protein, or in generating antisense DNA or RNA. The purified polypeptide
CC is useful for generating antibodies specific for it. The present sequence
CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20030073623
XX
XX Sequence 461 BP; 160 A; 64 C; 96 G; 126 T; 0 U; 15 Other;
SQ
Query Match 70.4%; Score 17.6; DB 9; Length 461;
Best Local Similarity 83.3%; Pred. No. 3.9e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 2 AAAAAGTCATCCCAATGTTATG 25
Db 209 AAAAAGTCATCAACACTGTTAAG 186
|||||
RESULT 34
AAH09944/c
ID AAH09944 standard; cdna; 549 BP.
XX
XX AAH09944;
XX
XX 26-JUN-2001 (first entry)
XX
XX Human cdna clone (3'-primer) SEQ ID NO:6779.
XX
XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
XX Homo sapiens.
XX
XX EP1074617-A2.
XX
XX 07-FEB-2001.
XX
XX 28-JUL-2000; 2000EP-00116126.
XX
XX 29-JUL-1999; 99JP-00248036.
XX
XX 27-AUG-1999; 99JP-00300253.
XX
XX 11-JAN-2000; 2000JP-00118776.
XX
XX 02-MAY-2000; 2000JP-00183767.
XX
XX 09-JUN-2000; 2000JP-00241899.
XX
XX (HELI-) HELIX RES INST.
XX
XX Ota T, Isogai T, Nishikawa T, Hayaashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
XX WPI; 2001-318749/34.
XX
```


PR 20-OCT-2000; 2000US-02411785P.
 PR 20-OCT-2000; 2000US-02411786P.
 PR 20-OCT-2000; 2000US-02411787P.
 PR 20-OCT-2000; 2000US-02411808P.
 PR 20-OCT-2000; 2000US-02411809P.
 PR 20-OCT-2000; 2000US-02411826P.
 PR 01-NOV-2000; 2000US-0244617P.
 PR 08-NOV-2000; 2000US-0246474P.
 PR 08-NOV-2000; 2000US-0246475P.
 PR 08-NOV-2000; 2000US-0246476P.
 PR 08-NOV-2000; 2000US-0246477P.
 PR 08-NOV-2000; 2000US-0246478P.
 PR 08-NOV-2000; 2000US-0246523P.
 PR 08-NOV-2000; 2000US-0246524P.
 PR 08-NOV-2000; 2000US-0246525P.
 PR 08-NOV-2000; 2000US-0246526P.
 PR 08-NOV-2000; 2000US-0246527P.
 PR 08-NOV-2000; 2000US-0246528P.
 PR 08-NOV-2000; 2000US-0246532P.
 PR 08-NOV-2000; 2000US-0246609P.
 PR 08-NOV-2000; 2000US-0246610P.
 PR 08-NOV-2000; 2000US-0246611P.
 PR 08-NOV-2000; 2000US-0246613P.
 PR 17-NOV-2000; 2000US-0249207P.
 PR 17-NOV-2000; 2000US-0249208P.
 PR 17-NOV-2000; 2000US-0249209P.
 PR 17-NOV-2000; 2000US-0249210P.
 PR 17-NOV-2000; 2000US-0249211P.
 PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249244P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249246P.
 PR 17-NOV-2000; 2000US-0249285P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 05-DEC-2000; 2000US-0256719P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 08-DEC-2000; 2000US-0251990P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Rosen CA, Barash SC, Ruben SM;
 XX
 XX WPT; 2001-476224/51.
 XX
 PT Isolated polypeptide for treating, preventing and/or prognosing
 PT disorders related to the respiratory system including respiratory cancers
 PT and also for testing and detection e.g. diagnosis.
 XX
 XX Disclosure; SED ID No 1132; 546pp; English.
 PS
 XX The present invention relates to the isolation of novel human respiratory
 CC antigens (AAU17695-AAU17975), and cDNA and genomic sequences encoding for
 CC these polypeptides. The sequences of the invention are useful for
 CC preventing, treating and/or prognosing disorders related to the
 CC respiratory system including throat disorders (e.g. vocal cord paralysis,

CC tonsillitis, and laryngitis), lung disorders e.g. pneumonia, allergic
 CC disorders e.g. asthma, pleurisy, cystic fibrosis, emphysema, nose
 CC disorders and cancers of the respiratory tissues e.g. lung cancer. The
 CC polynucleotide sequences of the invention are useful in gene therapy and
 CC antisense therapy. AAS28161-AAS28764 represent genomic sequences encoding
 CC for novel human respiratory antigens. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 XX Sequence 596 BP; 201 A; 78 C; 92 G; 225 T; 0 U; 0 Other;
 SQ
 Query Match 70.4%; Score 17.6; DB 4; Length 596;
 Best Local Similarity 83.3%; Pred. No. 4.1e+02;
 Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 2 AAAAAGTGCATCCCCCAATGTTATG 25
 DB 290 AAAAAGTGAATCCCCCAATGTTATG 267
 RESULT 36
 ADG41894/C
 ID ADG41894 standard; DNA; 596 BP.
 XX
 AC ADG41894;
 XX
 DT 26-FEB-2004 (first entry)
 XX
 DE Human respiratory system associated genomic DNA seq id 1132.
 XX
 KW antiinflammatory; antiallergic; antiasthmatic; cytostatic; gene therapy;
 KW human respiratory system antigen;
 KW human respiratory system associated polynucleotide;
 KW respiratory system disorder; throat disorder; vocal cord paralysis;
 KW tonsillitis; laryngitis; lung disorder; pneumonia; allergic disorder;
 KW asthma; eosinophilic pneumonia; pleurisy; cystic fibrosis; emphysema;
 KW histiocytosis; sarcoidosis; nose disorder; rhinitis; sinusitis; neoplasm;
 KW cancer; respiratory tissue cancer; throat cancer; lung cancer;
 KW cancer of the nose; gene therapy; chromosome identification; forensic;
 KW human respiratory system associated protein; ds; human.
 XX
 OS Homo sapiens.
 XX
 XX US2003215893-A1.
 XX
 XX 20-NOV-2003.
 XX
 XX 07-AUG-2002; 2002US-00212872.
 XX
 XX 31-JAN-2000; 2000US-0179065P.
 XX 04-FEB-2000; 2000US-0180628P.
 XX 24-FEB-2000; 2000US-0184664P.
 XX 02-MAR-2000; 2000US-0186350P.
 XX 16-MAR-2000; 2000US-0189874P.
 XX 17-MAR-2000; 2000US-0190076P.
 XX 18-APR-2000; 2000US-0198123P.
 XX 19-MAY-2000; 2000US-0205515P.
 XX 07-JUN-2000; 2000US-0209467P.
 XX 28-JUN-2000; 2000US-0214886P.
 XX 30-JUN-2000; 2000US-0215135P.
 XX 07-JUL-2000; 2000US-0216647P.
 XX 07-JUL-2000; 2000US-0216880P.
 XX 11-JUL-2000; 2000US-0217487P.
 XX 14-JUL-2000; 2000US-0217496P.
 XX 14-JUL-2000; 2000US-0218290P.
 XX 26-JUL-2000; 2000US-0220963P.
 XX 26-JUL-2000; 2000US-0220964P.
 XX 14-AUG-2000; 2000US-0224518P.
 XX 14-AUG-2000; 2000US-0224519P.
 XX 14-AUG-2000; 2000US-0225213P.
 XX 14-AUG-2000; 2000US-0225214P.
 XX 14-AUG-2000; 2000US-0225266P.

PR 14-AUG-2000; 2000US-02255267P.
 PR 14-AUG-2000; 2000US-02255268P.
 PR 14-AUG-2000; 2000US-02255270P.
 PR 14-AUG-2000; 2000US-02255477P.
 PR 14-AUG-2000; 2000US-02255757P.
 PR 14-AUG-2000; 2000US-02255758P.
 PR 14-AUG-2000; 2000US-02255759P.
 PR 14-AUG-2000; 2000US-02255799P.
 PR 14-AUG-2000; 2000US-0226279P.
 PR 22-AUG-2000; 2000US-0226681P.
 PR 22-AUG-2000; 2000US-0226868P.
 PR 22-AUG-2000; 2000US-0227182P.
 PR 23-AUG-2000; 2000US-0227009P.
 PR 30-AUG-2000; 2000US-0228924P.
 PR 01-SEP-2000; 2000US-0229287P.
 PR 01-SEP-2000; 2000US-0229343P.
 PR 01-SEP-2000; 2000US-0229344P.
 PR 01-SEP-2000; 2000US-0229345P.
 PR 05-SEP-2000; 2000US-0229509P.
 PR 05-SEP-2000; 2000US-0229513P.
 PR 06-SEP-2000; 2000US-0230437P.
 PR 06-SEP-2000; 2000US-0230438P.
 PR 08-SEP-2000; 2000US-0231124P.
 PR 08-SEP-2000; 2000US-02311243P.
 PR 08-SEP-2000; 2000US-02311244P.
 PR 08-SEP-2000; 2000US-0231113P.
 PR 08-SEP-2000; 2000US-0231141P.
 PR 08-SEP-2000; 2000US-0232080P.
 PR 08-SEP-2000; 2000US-0232081P.
 PR 12-SEP-2000; 2000US-0231368P.
 PR 14-SEP-2000; 2000US-0232397P.
 PR 14-SEP-2000; 2000US-0232398P.
 PR 14-SEP-2000; 2000US-0232399P.
 PR 14-SEP-2000; 2000US-0232400P.
 PR 14-SEP-2000; 2000US-0232401P.
 PR 14-SEP-2000; 2000US-0233063P.
 PR 14-SEP-2000; 2000US-0233064P.
 PR 14-SEP-2000; 2000US-0233065P.
 PR 21-SEP-2000; 2000US-0234223P.
 PR 21-SEP-2000; 2000US-0234274P.
 PR 25-SEP-2000; 2000US-0234997P.
 PR 25-SEP-2000; 2000US-0234998P.
 PR 26-SEP-2000; 2000US-0235484P.
 PR 27-SEP-2000; 2000US-0235834P.
 PR 27-SEP-2000; 2000US-0235836P.
 PR 28-SEP-2000; 2000US-0235935P.
 PR 29-SEP-2000; 2000US-0236327P.
 PR 29-SEP-2000; 2000US-0236367P.
 PR 29-SEP-2000; 2000US-0236368P.
 PR 29-SEP-2000; 2000US-0236369P.
 PR 29-SEP-2000; 2000US-0236370P.
 PR 02-OCT-2000; 2000US-0236802P.
 PR 02-OCT-2000; 2000US-0237037P.
 PR 02-OCT-2000; 2000US-0237038P.
 PR 02-OCT-2000; 2000US-0237039P.
 PR 02-OCT-2000; 2000US-0237040P.
 PR 13-OCT-2000; 2000US-0239337P.
 PR 20-OCT-2000; 2000US-0240960P.
 PR 20-OCT-2000; 2000US-0241221P.
 PR 20-OCT-2000; 2000US-0241785P.
 PR 20-OCT-2000; 2000US-0241786P.
 PR 20-OCT-2000; 2000US-0241787P.
 PR 20-OCT-2000; 2000US-0241808P.
 PR 20-OCT-2000; 2000US-0241809P.
 PR 20-OCT-2000; 2000US-0241826P.
 PR 01-NOV-2000; 2000US-0246178P.
 PR 08-NOV-2000; 2000US-0246474P.
 PR 08-NOV-2000; 2000US-0246475P.
 PR 08-NOV-2000; 2000US-0246476P.
 PR 08-NOV-2000; 2000US-0246477P.
 PR 08-NOV-2000; 2000US-0246478P.
 PR 08-NOV-2000; 2000US-0246523P.
 PR 08-NOV-2000; 2000US-0246524P.
 PR 08-NOV-2000; 2000US-0246525P.

PR 08-NOV-2000; 2000US-0246526P.
 PR 08-NOV-2000; 2000US-0246527P.
 PR 08-NOV-2000; 2000US-0246528P.
 PR 08-NOV-2000; 2000US-0246532P.
 PR 08-NOV-2000; 2000US-0246609P.
 PR 08-NOV-2000; 2000US-0246610P.
 PR 08-NOV-2000; 2000US-0246611P.
 PR 08-NOV-2000; 2000US-0246613P.
 PR 17-NOV-2000; 2000US-0249207P.
 PR 17-NOV-2000; 2000US-0249208P.
 PR 17-NOV-2000; 2000US-0249209P.
 PR 17-NOV-2000; 2000US-0249210P.
 PR 17-NOV-2000; 2000US-0249211P.
 PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249244P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249264P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 05-DEC-2000; 2000US-0256719P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 08-DEC-2000; 2000US-0251990P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.
 PR 17-JAN-2001; 2001US-00764860.
 PR 14-FEB-2002; 2002US-00074095.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Ruben SM, Barash SC;
 XX WPI; 2003-902033/82.
 XX
 DR Novel respiratory system antigen and polynucleotides encoding the
 PT polypeptides, useful for treating diagnosing, treating or preventing
 PT tonsillitis, pneumonia, asthma and cystic fibrosis, emphysema, throat
 cancer.
 XX
 PS Disclosure; SEQ ID NO 1132; 236pp; English.
 XX
 CC The invention describes an isolated polypeptide (I) comprising an amino
 CC acid sequence that is at least 90% identical to polypeptide fragment of
 CC any one of 299 respiratory system antigen sequences (PS) and having
 CC biological activity, polypeptide domain or epitope of PS, full-length
 CC protein of PS, or variant, allelic variant or species homolog of PS. (I)
 CC or a polynucleotide (II) encoding (I) is also useful for diagnosing a
 CC pathological condition or a susceptibility to a pathological condition in
 CC a subject which involves determining the presence or absence of mutation
 CC in (II) or determining the presence or amount of expression of (I) in a
 CC biological sample and diagnosing a pathological condition based on the
 CC result. The human respiratory system associated polynucleotides, the
 CC polypeptides encoded by them, and antibodies that immunospecifically bind
 CC these polypeptides are useful in diagnosis, treatment, prevention and/or
 CC prognosis of disorders of respiratory system such as throat disorders
 CC (e.g., vocal cord paralysis, tonsillitis, and laryngitis), lung disorders
 CC (e.g., pneumonia), allergic disorders, (e.g., asthma and eosinophilic
 CC pneumonia), pleurisy, cystic fibrosis, emphysema, histiocytosis,

CC sarcoidosis, nose disorders (rhinitis and sinusitis), neoplasms and/or
CC cancers of respiratory tissues (e.g., throat cancer, lung cancer, and
CC cancer of the nose). The polynucleotides are useful in gene therapy
CC techniques, for chromosome identification, identifying individuals from

Query Match 70.4%; Score 17.6; DB 10; Length 596;

Best Local Similarity 83.3%; Pred. No. 4.1e+02;

Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AAAAAGTCATCCCATGTTATG 25

Db 290 AAAAAGTCATCCCATGTTATG 267

RESULT 37

AD197668/C

ID AD197668 standard; DNA; 596 BP.

XX

XX

AC

AD197668;

XX

DT

04-NOV-2004

(first entry)

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
PR 17-JAN-2001; 2001US-00764860.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Ruben SM, Barash SC;
PI WPI; 2003-765403/72.
XX
XX
XX New human respiratory system-related polypeptide and genes, useful for
PT treating, preventing or diagnosing e.g. pneumonia, lung cancer, cystic
PT fibrosis, asthma, sarcoidosis, rhinitis, leukemia, inflammations or
PT sinusitis.
XX
XX Disclosure; SEQ ID NO 1132; 202pp; English.
XX
XX This invention is related to a novel isolated polypeptide, which
CC comprises a human respiratory system-related polypeptide, and the DNA
CC sequence which encodes it. The invention may be useful for the
CC development of compounds with an antiasthmatic, antibacterial,
CC antiinflammatory, cytostatic, antianaemic or antiallergic activity. In
CC addition, the sequences disclosed may be useful for gene therapy. The
CC polypeptide or polynucleotide is useful for treating, preventing or
CC ameliorating a medical condition, for example pneumonia, lung cancer,
CC cystic fibrosis, asthma, sarcoidosis, rhinitis, anaemia, leukaemia,
CC inflammations, sinusitis, chronic obstructive pulmonary disease or
CC infectious diseases. The polypeptide or polynucleotide is also useful for
CC diagnosing any of these diseases or a susceptibility to the disease. The
CC present sequence is that of a human DNA sequence which is related to a
CC human respiratory system associated gene of the invention.
XX
SQ Sequence 596 BP; 201 A; 78 C; 92 G; 225 T; 0 U; 0 Other;

Query Match 70.4%; Score 17.6; DB 11; Length 596;
Best Local Similarity 83.3%; Pred. No. 4.1e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AAAAAGTGCATCCCAATGTTATG 25
|||||
Db 290 AAAAAGTGCATCCCAATGTTATG 267

RESULT 38
ABV56236/c
ID ABV56236 standard; cDNA; 612 BP.
XX

AC ABV56236;
XX
DT 17-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 56227.
XX
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200160860-A2.
XX
PD 23-AUG-2001.
XX
PF 20-FEB-2001; 2001WO-US005171.
XX
PR 17-FEB-2000; 2000US-0183319P.
PR 16-MAR-2000; 2000US-0189862P.
PR 25-MAY-2000; 2000US-0207454P.
PR 09-JUN-2000; 2000US-0211314P.
PR 18-JUL-2000; 2000US-0219007P.
PR 13-DEC-2000; 2000US-0255281P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Schlegel R, Endege WO, Monahan JE;
PI WPI; 2001-662795/76.
XX
XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX
PS Claim 1; Page 10854; 11750pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (i) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
SQ Sequence 612 BP; 173 A; 134 C; 140 G; 165 T; 0 U; 0 Other;

Query Match 70.4%; Score 17.6; DB 5; Length 612;
Best Local Similarity 83.3%; Pred. No. 4.1e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AAAAAGTGCATCCCAATGTTATG 25
|||||
Db 276 AAAAAGTGCATCCCAATGTTATG 253

RESULT 39
AAS28126
ID AAS28126 standard; cDNA; 614 BP.
XX
AC AAS28126;
XX
DT 07-NOV-2001 (first entry)
XX
DE Novel cDNA encoding for human respiratory antigen #258.
XX
KW Human; respiratory antigen; respiratory disorder; throat disorder;
KW lung disorder; nose disorder; lung cancer; gene therapy; cytostatic;
KW anti allergic; anti asthmatic; anti inflammatory; olfactory;

KW respiratory active; ss.
XX Homo sapiens.
OS WO200155448-A1.
PN 02-AUG-2001.
PD 17-JAN-2001; 2001WO-US001333.
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 08-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234937P.
PR 25-SEP-2000; 2000US-0234938P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.

PR 26-JUL-2000; 2000US-0220963P.
 PR 26-JUL-2000; 2000US-0220964P.
 PR 14-AUG-2000; 2000US-0224518P.
 PR 14-AUG-2000; 2000US-0224519P.
 PR 14-AUG-2000; 2000US-0225213P.
 PR 14-AUG-2000; 2000US-0225214P.
 PR 14-AUG-2000; 2000US-0225266P.
 PR 14-AUG-2000; 2000US-0225267P.
 PR 14-AUG-2000; 2000US-0225268P.
 PR 14-AUG-2000; 2000US-0225270P.
 PR 14-AUG-2000; 2000US-0225447P.
 PR 14-AUG-2000; 2000US-0225757P.
 PR 14-AUG-2000; 2000US-0225758P.
 PR 14-AUG-2000; 2000US-0225759P.
 PR 18-AUG-2000; 2000US-0226279P.
 PR 22-AUG-2000; 2000US-0226681P.
 PR 22-AUG-2000; 2000US-0226868P.
 PR 22-AUG-2000; 2000US-0227182P.
 PR 23-AUG-2000; 2000US-0227009P.
 PR 30-AUG-2000; 2000US-0228924P.
 PR 01-SEP-2000; 2000US-0229287P.
 PR 01-SEP-2000; 2000US-0229343P.
 PR 01-SEP-2000; 2000US-0229344P.
 PR 01-SEP-2000; 2000US-0229345P.
 PR 05-SEP-2000; 2000US-0229509P.
 PR 05-SEP-2000; 2000US-0229513P.
 PR 06-SEP-2000; 2000US-0230437P.
 PR 06-SEP-2000; 2000US-0230438P.
 PR 08-SEP-2000; 2000US-0231242P.
 PR 08-SEP-2000; 2000US-0231243P.
 PR 08-SEP-2000; 2000US-0231244P.
 PR 08-SEP-2000; 2000US-0231413P.
 PR 08-SEP-2000; 2000US-0231414P.
 PR 08-SEP-2000; 2000US-0232080P.
 PR 08-SEP-2000; 2000US-0232081P.
 PR 12-SEP-2000; 2000US-0231968P.
 PR 14-SEP-2000; 2000US-0232397P.
 PR 14-SEP-2000; 2000US-0232398P.
 PR 14-SEP-2000; 2000US-0232399P.
 PR 14-SEP-2000; 2000US-0232400P.
 PR 14-SEP-2000; 2000US-0232401P.
 PR 14-SEP-2000; 2000US-0232403P.
 PR 14-SEP-2000; 2000US-0232404P.
 PR 21-SEP-2000; 2000US-0234223P.
 PR 21-SEP-2000; 2000US-0234274P.
 PR 25-SEP-2000; 2000US-0234997P.
 PR 25-SEP-2000; 2000US-0234998P.
 PR 26-SEP-2000; 2000US-0235484P.
 PR 27-SEP-2000; 2000US-0235834P.
 PR 27-SEP-2000; 2000US-0235836P.
 PR 29-SEP-2000; 2000US-0236327P.
 PR 29-SEP-2000; 2000US-0236367P.
 PR 29-SEP-2000; 2000US-0236368P.
 PR 29-SEP-2000; 2000US-0236369P.
 PR 29-SEP-2000; 2000US-0236370P.
 PR 02-OCT-2000; 2000US-0236802P.
 PR 02-OCT-2000; 2000US-0237037P.
 PR 02-OCT-2000; 2000US-0237038P.
 PR 02-OCT-2000; 2000US-0237039P.
 PR 02-OCT-2000; 2000US-0237040P.
 PR 13-OCT-2000; 2000US-0239335P.
 PR 13-OCT-2000; 2000US-0239937P.
 PR 20-OCT-2000; 2000US-0240960P.
 PR 20-OCT-2000; 2000US-0241221P.
 PR 20-OCT-2000; 2000US-0241785P.
 PR 20-OCT-2000; 2000US-0241786P.
 PR 20-OCT-2000; 2000US-0241787P.
 PR 20-OCT-2000; 2000US-0241808P.
 PR 20-OCT-2000; 2000US-0241809P.
 PR 20-OCT-2000; 2000US-0241826P.
 PR 01-NOV-2000; 2000US-0244617P.
 PR 08-NOV-2000; 2000US-0246474P.

PR 08-NOV-2000; 2000US-0246475P.
 PR 08-NOV-2000; 2000US-0246476P.
 PR 08-NOV-2000; 2000US-0246477P.
 PR 08-NOV-2000; 2000US-0246478P.
 PR 08-NOV-2000; 2000US-0246523P.
 PR 08-NOV-2000; 2000US-0246524P.
 PR 08-NOV-2000; 2000US-0246525P.
 PR 08-NOV-2000; 2000US-0246526P.
 PR 08-NOV-2000; 2000US-0246527P.
 PR 08-NOV-2000; 2000US-0246528P.
 PR 08-NOV-2000; 2000US-0246532P.
 PR 08-NOV-2000; 2000US-0246609P.
 PR 08-NOV-2000; 2000US-0246610P.
 PR 08-NOV-2000; 2000US-0246611P.
 PR 08-NOV-2000; 2000US-0246613P.
 PR 17-NOV-2000; 2000US-0249207P.
 PR 17-NOV-2000; 2000US-0249208P.
 PR 17-NOV-2000; 2000US-0249209P.
 PR 17-NOV-2000; 2000US-0249210P.
 PR 17-NOV-2000; 2000US-0249211P.
 PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249264P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 05-DEC-2000; 2000US-0255719P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 08-DEC-2000; 2000US-0251990P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.
 PR 17-JAN-2001; 2001US-00764860.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Ruben SM, Barash SC;

WPI; 2003-765403/72.

P-PSDB; ADI97096.

New human respiratory system-related polypeptide and genes, useful for treating, preventing or diagnosing e.g. pneumonia, lung cancer, cystic fibrosis, asthma, sarcoidosis, rhinitis, leukemia, inflammations or sinusitis.

Claim 1; SEQ ID NO 268; 202pp; English.

This invention is related to a novel isolated polypeptide, which comprises a human respiratory system-related polypeptide, and the DNA sequence which encodes it. The invention may be useful for the development of compounds with an antiasthmatic, antibacterial, antiinflammatory, cytostatic, antianaemic or antiallergic activity. In addition, the sequences disclosed may be useful for gene therapy. The polypeptide or polynucleotide is useful for treating, preventing or ameliorating a medical condition, for example pneumonia, lung cancer, cystic fibrosis, asthma, sarcoidosis, rhinitis, leukemia, leukaemia, inflammations, sinusitis, chronic obstructive pulmonary disease or

CC infectious diseases. The polypeptide or polynucleotide is also useful for
CC diagnosing any of these diseases or a susceptibility to the disease. The
CC present sequence is that of a respiratory system associated human gene of
CC the invention.

XX Sequence 614 BP; 231 A; 96 C; 80 G; 207 T; 0 U; 0 Other;

Query Match 70.4%; Score 17.6; DB 11; Length 614;

Best Local Similarity 83.3%; Pred. No. 4.1e+02;

Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AAAAAGTCATCCCAATGTTAT 25

DB 309 AAAAAGTCATCCCAATGTTAT 332

RESULT 42

ACA53946

ID ACA53946 standard; DNA; 1086 BP.

XX ACA53946;

AC ACA53946;

XX 19-JUN-2003 (first entry)

XX Prokaryotic essential gene #35603.

XX Antisense; ds; prokaryotic essential gene; cell proliferation;

XX drug design; gene.

XX Versinia pestis.

XX WO200277183-A2.

XX 03-OCT-2002.

XX 21-MAR-2002; 2002WO-US009107.

XX 21-MAR-2001; 2001US-00815242.

XX 06-SEP-2001; 2001US-00948993.

XX 25-OCT-2001; 2001US-0342923P.

XX 08-FEB-2002; 2002US-00072851.

XX 06-MAR-2002; 2002US-0362699P.

XX (ELIT-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;

XX Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX WPI; 2003-029926/02.

XX P-PSDB; ABUS0076.

XX New antisense nucleic acids, useful for identifying proteins or screening
XX for homologous nucleic acids required for cellular proliferation to
XX isolate candidate molecules for rational drug discovery programs.

XX Claim 14; SEQ ID NO 41816; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of
XX the 6213 antisense sequences given in the specification where expression
XX of the nucleic acid inhibits proliferation of a cell. Also included are:
XX (1) a vector comprising a promoter operably linked to the nucleic acid
XX encoding a polypeptide whose expression is inhibited by the antisense
XX nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX polypeptide or its fragment whose expression is inhibited by the
XX antisense nucleic acid; (4) an antibody capable of specifically binding
XX the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
XX proliferation or the activity of a gene in an operon required for
XX proliferation; (7) identifying a compound that influences the activity of
XX the gene product or that has an activity against a biological pathway
XX required for proliferation, or that inhibits cellular proliferation; (8)
XX identifying a gene required for cellular proliferation or the biological
XX pathway in which a proliferation-required gene or its gene product lies
XX or a gene on which the test compound that inhibits proliferation of an

CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
CC prokaryotic essential genes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 1086 BP; 283 A; 233 C; 282 G; 288 T; 0 U; 0 Other;

Query Match 70.4%; Score 17.6; DB 8; Length 1086;

Best Local Similarity 83.3%; Pred. No. 4.4e+02;

Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAGTCATCCCAATGTTAT 24

DB 333 AAAAAGTCATCCCAATGTTAT 356

RESULT 43

ACA43061

ID ACA43061 standard; DNA; 1506 BP.

XX ACA43061;

AC ACA43061;

XX 19-JUN-2003 (first entry)

XX Prokaryotic essential gene #24718.

XX Antisense; ds; prokaryotic essential gene; cell proliferation;

XX drug design; gene.

XX Pasteurella multocida.

XX WO200277183-A2.

XX 03-OCT-2002.

XX 21-MAR-2002; 2002WO-US009107.

XX 21-MAR-2001; 2001US-00815242.

XX 06-SEP-2001; 2001US-00948993.

XX 25-OCT-2001; 2001US-0342923P.

XX 08-FEB-2002; 2002US-00072851.

XX 06-MAR-2002; 2002US-0362699P.

XX (ELIT-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;

XX Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX WPI; 2003-029926/02.

XX P-PSDB; ABUS39191.

XX New antisense nucleic acids, useful for identifying proteins or screening
XX for homologous nucleic acids required for cellular proliferation to
XX isolate candidate molecules for rational drug discovery programs.

XX Claim 14; SEQ ID NO 30931; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of
XX the 6213 antisense sequences given in the specification where expression
XX of the nucleic acid inhibits proliferation of a cell. Also included are:
XX (1) a vector comprising a promoter operably linked to the nucleic acid
XX encoding a polypeptide whose expression is inhibited by the antisense

CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
CC prokaryotic essential genes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 1506 BP; 472 A; 316 C; 304 G; 414 T; 0 U; 0 Other;

Query Match 70.4%; Score 17.6; DB 8; Length 1506;
Best Local Similarity 83.3%; Pred. No. 4.6e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AAAAAGTCATCCCAATGTTATG 25
DB 263 AAAAAGTCATCCCAATGTTATG 286

RESULT 44
ACH94314/c
ID ACH94314 standard; DNA; 1770 BP.

AC ACH94314;

XX 29-JUL-2004 (first entry)

XX Klebsiella pneumoniae polynucleotide seqid 109.

XX Recombinant expression vector; transcription regulatory element;
XX Klebsiella pneumoniae protein; antibacterial; Vaccine; gene; ds.

XX Klebsiella pneumoniae.

XX US6610836-B1.

XX 26-AUG-2003.

XX 27-JAN-2000; 2000US-00489039.

XX 29-JAN-1999; 99US-0117747P.

XX (GENO-) GENOME THERAPEUTICS CORP.

XX Breton GL, Osborne M;

XX WPI; 2003-895346/82.

XX P-PSDB; ABO60763.

XX New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for
PT preparing a vaccine composition against Klebsiella pneumoniae.

XX Disclosure; SEQ ID NO 109; 932pp; English.

CC The invention describes a new isolated nucleic acid encoding a Klebsiella
CC pneumoniae polypeptide. Also described are: a recombinant expression
CC vector comprising the nucleic acid, operably linked to a transcription
CC regulatory element; and a cell comprising the recombinant expression
CC vector. The nucleic acid is useful for preparing a vaccine composition
CC against Klebsiella pneumoniae. This sequence encodes a Klebsiella
CC pneumoniae polypeptide of the invention

XX Sequence 1770 BP; 305 A; 509 C; 569 G; 387 T; 0 U; 0 Other;

Query Match 70.4%; Score 17.6; DB 11; Length 1770;
Best Local Similarity 83.3%; Pred. No. 4.7e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AAAAAGTCATCCCAATGTTATG 25

DB 1338 AAAAAGTCATCCCAATGTTATG 1315

RESULT 45

ACH94414

ID ACH94414 standard; DNA; 1803 BP.

XX ACH94414;

XX 29-JUL-2004 (first entry)

XX Klebsiella pneumoniae polynucleotide seqid 209.

XX Recombinant expression vector; transcription regulatory element;
XX Klebsiella pneumoniae protein; antibacterial; Vaccine; gene; ds.

XX Klebsiella pneumoniae.

XX US6610836-B1.

XX 26-AUG-2003.

XX 27-JAN-2000; 2000US-00489039.

XX 29-JAN-1999; 99US-0117747P.

XX (GENO-) GENOME THERAPEUTICS CORP.

XX Breton GL, Osborne M;

XX WPI; 2003-895346/82.

XX P-PSDB; ABO60863.

XX New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for
PT preparing a vaccine composition against Klebsiella pneumoniae.

XX Disclosure; SEQ ID NO 209; 932pp; English.

XX The invention describes a new isolated nucleic acid encoding a Klebsiella
CC pneumoniae polypeptide. Also described are: a recombinant expression
CC vector comprising the nucleic acid, operably linked to a transcription
CC regulatory element; and a cell comprising the recombinant expression
CC vector. The nucleic acid is useful for preparing a vaccine composition
CC against Klebsiella pneumoniae. This sequence encodes a Klebsiella
CC pneumoniae polypeptide of the invention

XX Sequence 1803 BP; 397 A; 573 C; 524 G; 309 T; 0 U; 0 Other;

Query Match 70.4%; Score 17.6; DB 11; Length 1803;
Best Local Similarity 83.3%; Pred. No. 4.8e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AAAAAGTCATCCCAATGTTATG 25

DB 406 AAAAAGTCATCCCAATGTTATG 429


```
RESULT 46
ADR60058
ID ADR60058 standard; cDNA; 1951 BP.
XX
AC ADR60058;
AD 02-DEC-2004 (first entry)
DT
DE Cotton cDNA sequence, SEQ ID 839.
XX
KW Cotton; ss; plant; cold tolerance; growth rate; cell cycle pathway;
KW drought tolerance; plant disease resistance; galactomannan; lignin;
KW plant growth regulator; heat tolerance; herbicide tolerance;
KW homologous recombination; extreme osmotic condition tolerance;
KW pathogen resistance; pest resistance; yield; photosynthesis; seed oil;
KW stress resistance.
XX
OS Gossypium hirsutum.
XX
PN US2004181830-A1.
XX
PD 16-SEP-2004.
XX
PF 29-JAN-2004; 2004US-00767795.
XX
PR 07-MAY-2001; 2001US-00849529.
PR 12-DEC-2001; 2001US-00021323.
XX
PA (KOVA/) KOVALIC D K.
PA (ZHOU/) ZHOU Y.
PA (CAOV/) CAO Y.
PI Kovalic DK, Zhou Y, Cao Y;
XX
WPI; 2004-667718/65.
XX
XX New recombinant nucleic acid molecules and polypeptides from Gossypium
PT hirsutum, useful for producing plants with improved biological
PT characteristics (e.g. improved plant cold or drought tolerance).
XX
Claim 1; SEQ ID NO 839; 14pp; English.
XX
CC The invention relates to a recombinant polynucleotide comprising any of
CC the 58798 Cotton plant cDNA sequences mentioned in the specification.
CC Also a recombinant polypeptide comprising any of the 58798 amino acid
CC sequences mentioned in the specification and producing a plant having an
CC improved property. Producing a plant having an improved property
CC comprises transforming a plant with a recombinant construct comprising a
CC promoter region functional in a plant cell operably joined to a
CC polynucleotide comprising a coding sequence for a polypeptide associated
CC with the property, and growing the transformed plant. The polypeptide is
CC useful for improving plant cold tolerance, manipulating growth rate in
CC plant cells by modification of the cell cycle pathway, improving plant
CC drought tolerance, providing increased resistance to plant disease,
CC producing galactomannan (or lignin or plant growth regulators), improving
CC plant heat tolerance, improving plant tolerance to herbicides, increasing
CC the rate of homologous recombination in plants, improving plant tolerance
CC to extreme osmotic conditions or to pathogens or pests, improving yield
CC by modification of photosynthesis, modifying seed oil or protein yield
CC and/or content, improving yield by modification of carbohydrate, nitrogen
CC or phosphorus use and/or uptake, or improving yield by providing improved
CC plant growth and development under at least one stress condition. The
CC polynucleotide and polypeptide may also be used in recombinant DNA
CC constructs, in physical arrays of molecules, as plant breeding markers,
CC or in computer-based storage and analysis systems. The present sequence
CC is a Cotton plant cDNA of the invention. NOTE: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?docID=20040181830. However only 6585
CC polynucleotide sequences were available, the remaining 52213
CC polynucleotides and all 58798 protein sequences were not present.
XX
SQ Sequence 1951 BP; 519 A; 417 C; 395 G; 620 T; 0 U; 0 Other;

Query Match 70.4%; Score 17.6; DB 13; Length 1951;
Best Local Similarity 83.3%; Pred. No. 4.8e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 2 AAAAACTGCATCCCAATGTTATG 25
DB 24 AAAAACTGCATCCCAATGTTTG 47

RESULT 47
ADB68941
ID ADB68941 standard; DNA; 2975 BP.
XX
AC ADB68941;
AD 04-DEC-2003 (first entry)
DT
DE C. neoformans genomic DNA sequence SEQ ID NO:68.
XX
KW ds; gene; fungicide; gene therapy; infection.
XX
OS Cryptococcus neoformans.
XX
PN WO2003052076-A2.
XX
PD 26-JUN-2003.
XX
PF 17-DEC-2002; 2002WO-US040225.
XX
PR 17-DEC-2001; 2001US-0341261P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Zamudio C, Eroshkin AM;
XX
WPI; 2003-533017/50.
XX
P-PSDB; ADB70024.
XX
XX New nucleic acid, useful for preparing a composition for treating an
PT infection caused by Cryptococcus neoformans.
XX
Claim 3; SEQ ID NO 68; 136pp; English.
XX
CC The invention relates to a novel purified or isolated Cryptococcus
CC neoformans nucleic acid molecule comprising a sequence encoding a
CC polypeptide comprising a sequence not given in the specification. A
CC polynucleotide of the invention has fungicide activity, and may have a
CC use in gene therapy. The nucleic acid is useful for preparing a
CC composition for treating an infection caused by Cryptococcus neoformans.
CC The present sequence represents a C. neoformans sequence of the
CC invention. Note: The sequence data for this patent is not represented in
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pat_sequences.
XX
SQ Sequence 2975 BP; 747 A; 739 C; 716 G; 772 T; 0 U; 1 Other;

Query Match 70.4%; Score 17.6; DB 10; Length 2975;
Best Local Similarity 83.3%; Pred. No. 5.1e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 AAAAACTGCATCCCAATGTTAT 24
DB 1974 AGAARACGGCATCGACATGTTAT 1997

RESULT 48
ADE25037/c
ID ADE25037 standard; cDNA; 3833 BP.
XX
AC ADE25037;
AD 29-JAN-2004 (first entry)
DT
```

XX DE Plant growth associated gene cDNA seq id 12.

XX KW plant growth; plant growth trait modulation; Brassicaceae; Arabidopsis;

XX KW Brassica; Zea; Oryza; Triticum; Hordeum; Lolium; Sorghum; Glycine;

XX KW Medicago; Helianthus; Lactuca; Beta; Vitis; Solanum; Lycopersicon;

XX KW Capsicum; Gossypium; Hevea; Linum; Prunus; Citrus; Populus; Pinus;

XX KW Quercus; gene; ss.

XX OS Magnoliophyta.

XX PN US2003188343-A1.

XX PD 02-OCT-2003.

XX PF 07-JAN-2003; 2003US-00338777.

XX PR 09-JAN-2002; 2002US-0347288P.

XX PX (LYNX-) LYNX THERAPEUTICS INC.

XX PI Bowen BA, Haudenschild CD, Buckler ES;

XX PX WPI: 2003-803305/75.

XX DR P-PSDB; ADE25067.

XX XX New isolated or recombinant polypeptide for use in modulating a plant

PT growth trait in a flowering plant e.g. in Arabidopsis, Brassica, Zea, or

PT Oryza.

XX PS Claim 1; SEQ ID NO 12; 81pp; English.

XX CC The invention describes an isolated or recombinant polypeptide (I)

CC comprising a sequence: (a) comprising 1 of 30 sequences (S1), as given in

CC the specification, or a conservative variant; (b) encoded by 1 of 30

CC sequences (S2), as given in the specification, or a conservative variant;

CC (c) encoded by a sequence that hybridises under stringent conditions to

CC S2; and (d) encoded by a sequence 70 % identical to S2. The expression or

CC activity of (I) is modulated to modulate a plant growth trait in a

CC flowering plant, of the family Brassicaceae, preferably in a plant that

CC is Arabidopsis, Brassica, Zea, Oryza, Triticum, Hordeum, Lolium, Sorghum,

CC Glycine, Medicago, Helianthus, Lactuca, Beta, Vitis, Solanum,

CC Lycopersicon, Capsicum, Gossypium, Hevea, Linum, Prunus, Citrus, Populus,

CC Pinus, or Quercus. A new method is used to detect genes for a plant

CC growth trait. This sequence represents a plant growth associated cDNA.

XX SQ Sequence 3833 BP; 1023 A; 705 C; 831 G; 1274 T; 0 U; 0 Other;

Query Match 70.4%; Score 17.6; DB 10; Length 3833;

Best Local Similarity 83.3%; Pred. No. 5.3e+02;

Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 AAAAACTGCATCCCAATGTTAT 24

||||| ||||| ||||| |||||

Db 1923 AAAAAATTCATCTCTCAATGTTAT 1900

RESULT 49

AAK68667/c

ID AAK68667 standard; DNA; 4963 BP.

XX AC AAK68667;

XX XX 06-NOV-2001 (first entry)

DT Human immune/haematopoietic antigen genomic sequence SEQ ID NO:23479.

DE Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;

XX KW cytostatic; gene therapy; vaccine; metastasis; ds.

XX OS Homo sapiens.

XX PN WO200157182-A2.

XX PD 09-AUG-2001.

XX PF 17-JAN-2001; 2001WO-US001354.

XX 31-JAN-2000; 2000US-0179065P.

PR 04-FEB-2000; 2000US-0180628P.

PR 24-FEB-2000; 2000US-0184664P.

PR 02-MAR-2000; 2000US-0186350P.

PR 16-MAR-2000; 2000US-0189874P.

PR 17-MAR-2000; 2000US-0190076P.

PR 18-APR-2000; 2000US-0198123P.

PR 19-MAY-2000; 2000US-0205515P.

PR 07-JUN-2000; 2000US-0209467P.

PR 28-JUN-2000; 2000US-0214886P.

PR 30-JUN-2000; 2000US-0215135P.

PR 07-JUL-2000; 2000US-0216647P.

PR 07-JUL-2000; 2000US-0216880P.

PR 11-JUL-2000; 2000US-0217487P.

PR 11-JUL-2000; 2000US-0217496P.

PR 14-JUL-2000; 2000US-0218290P.

PR 26-JUL-2000; 2000US-0220963P.

PR 26-JUL-2000; 2000US-0220964P.

PR 14-AUG-2000; 2000US-0224518P.

PR 14-AUG-2000; 2000US-0224519P.

PR 14-AUG-2000; 2000US-0225213P.

PR 14-AUG-2000; 2000US-0225214P.

PR 14-AUG-2000; 2000US-0225266P.

PR 14-AUG-2000; 2000US-0225267P.

PR 14-AUG-2000; 2000US-0225268P.

PR 14-AUG-2000; 2000US-0225270P.

PR 14-AUG-2000; 2000US-0225447P.

PR 14-AUG-2000; 2000US-0225757P.

PR 14-AUG-2000; 2000US-0225758P.

PR 14-AUG-2000; 2000US-0225759P.

PR 18-AUG-2000; 2000US-0226279P.

PR 22-AUG-2000; 2000US-0226681P.

PR 22-AUG-2000; 2000US-0226688P.

PR 22-AUG-2000; 2000US-0227182P.

PR 23-AUG-2000; 2000US-0227009P.

PR 30-AUG-2000; 2000US-0228924P.

PR 01-SEP-2000; 2000US-0229287P.

PR 01-SEP-2000; 2000US-0229343P.

PR 01-SEP-2000; 2000US-0229344P.

PR 01-SEP-2000; 2000US-0229345P.

PR 05-SEP-2000; 2000US-0229509P.

PR 05-SEP-2000; 2000US-0229513P.

PR 06-SEP-2000; 2000US-0230437P.

PR 06-SEP-2000; 2000US-0230438P.

PR 08-SEP-2000; 2000US-0231242P.

PR 08-SEP-2000; 2000US-0231243P.

PR 08-SEP-2000; 2000US-0231244P.

PR 08-SEP-2000; 2000US-0231413P.

PR 08-SEP-2000; 2000US-0231414P.

PR 08-SEP-2000; 2000US-0232080P.

PR 08-SEP-2000; 2000US-0232081P.

PR 12-SEP-2000; 2000US-0231968P.

PR 14-SEP-2000; 2000US-0232397P.

PR 14-SEP-2000; 2000US-0232398P.

PR 14-SEP-2000; 2000US-0232399P.

PR 14-SEP-2000; 2000US-0232400P.

PR 14-SEP-2000; 2000US-0232401P.

PR 14-SEP-2000; 2000US-0233063P.

PR 14-SEP-2000; 2000US-0233064P.

PR 14-SEP-2000; 2000US-0233065P.

PR 21-SEP-2000; 2000US-0234223P.

PR 21-SEP-2000; 2000US-0234274P.

PR 25-SEP-2000; 2000US-0234997P.

PR 25-SEP-2000; 2000US-0234998P.

PR 26-SEP-2000; 2000US-0235484P.

PR 27-SEP-2000; 2000US-0235834P.

PR 27-SEP-2000; 2000US-0235836P.

PR 29-SEP-2000; 2000US-0236327P.

PR 07-JUL-2000; 2000US-0215880P.
PR 11-JUL-2000; 2000US-02117487P.
PR 11-JUL-2000; 2000US-02117496P.
PR 14-JUL-2000; 2000US-02118290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0228287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234224P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236328P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239335P.
PR 13-OCT-2000; 2000US-0239337P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.

PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241825P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249298P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251473P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX
XX
PI Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-451937/48.
XX
XX
PT Isolated polypeptide for treating, preventing and/ or prognosing
PT disorders related to the musculoskeletal system including musculoskeletal
PT cancers and also for testing and detection e.g. diagnosis.
XX
XX
PS Example 2; SEQ ID NO 2903; 781pp + Sequence Listing; English.
XX
XX
CC The invention relates to novel genes (AAL34669-AAL37666) and proteins
CC (ABB03087-ABB04109) associated with the musculoskeletal system useful for
CC preventing, treating or ameliorating medical conditions e.g. by protein
CC or gene therapy. The genes are isolated from a range of human tissues
CC disclosed in the specification. The nucleic acids, proteins, antibodies
CC and (ant)agonists are useful in the diagnosis, treatment and prevention
CC of: (a) cancer, e.g. breast and ovarian cancer and other cancers of the
CC adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver,
CC lung, or urogenital; (b) immune disorders e.g. Addison's disease,
CC

CC allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
CC arthritis and ulcerative colitis; (c) cardiovascular disorders such as
CC myocardial ischaemias; (d) wound healing; (e) neurological diseases e.g.
CC cerebral anoxia and epilepsy; and (f) infectious diseases such as viral,
CC bacterial, fungal and parasitic infections. Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX
SQ Sequence 4963 BP; 1480 A; 991 C; 1089 G; 1403 T; 0 U; 0 Other;
Query Match 70.4%; Score 17.6; DB 4; Length 4963;
Best Local Similarity 83.3%; Pred. No. 5.5e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAACTGCATCCCAATGTTAT 24
||||| ||||| ||||| |||||
Db 2660 AAAAACTGGATTCCCAAGATAT 2637

Search completed: February 3, 2006, 21:57:05
Job time : 210.111 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 3, 2006, 14:38:25 ; Search time 2951.67 Seconds
(without alignments)
396.277 Million cell updates/sec

Title: US-10-719-900-25

Perfect score: 25

Sequence: 1 aaaaactgcatcccaatgttatg 25

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database :

EST:*

- 1: gb_est1.*
- 2: gb_est2.*
- 3: gb_est3.*
- 4: gb_hcc.*
- 5: gb_est4.*
- 6: gb_est5.*
- 7: gb_est6.*
- 8: gb_est7.*
- 9: gb_gss1.*
- 10: gb_gss2.*
- 11: gb_gss3.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	23.4	93.6	226	1	BB127535
2	23.4	93.6	274	1	AV259784
3	23.4	93.6	276	1	AV332113
C 4	23.4	93.6	340	6	CV76103
5	23.4	93.6	413	5	DY619222
C 6	23.4	93.6	419	1	AI847234
7	23.4	93.6	438	5	BY616767
8	23.4	93.6	449	2	BB750064
C 9	23.4	93.6	474	1	AI844638
10	23.4	93.6	648	1	BB130746
11	23.4	93.6	887	4	AK079373
12	22.4	89.6	175	1	BB083460
C 13	22.4	89.6	220	1	AI412710
14	22.4	89.6	279	2	BB347350
15	22.4	89.6	329	2	BB391724
C 16	22.4	89.6	369	2	BP404880
C 17	22.4	89.6	380	1	AA997380
C 18	21.8	87.2	512	2	BQ177061
19	21.8	87.2	647	2	BB629036
C 20	21.8	87.2	744	7	CO873358
21	21.8	87.2	764	7	CO878266
22	21	84.0	407	2	BG727345

QY 1 AAAAACTGCATCCCAATGTTATG 25
 |||||
 150 AAAAACTGCATGCCCAATGTTATG 174

RESULT 2
 AV259784 274 bp mRNA linear EST 04-NOV-1999
 LOCUS AV259784 RIKEN full-length enriched, adult male testis (DH10B) Mus
 DEFINITION musculus cDNA clone 4930406D09 3' similar to X62840 R.rattus mRNA
 for potassium channel protein (3145 bp), mRNA sequence.

ACCESSION AV259784
 VERSION AV259784.1 GI:6247243
 KEYWORDS Mus musculus (house mouse)
 SOURCE Mus musculus
 ORGANISM Mus musculus

REFERENCE
 AUTHORS Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 274)
 Konno, H., Aizawa, K., Akahira, S., Akiyama, J., Carninci, P., Endo, T., Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N., Hirozane, T., Hori, F., Ishii, Y., Ishikawa, T., Itoh, M., Izawa, M., Kadota, K., Kagawa, I., Kai, C., Kawai, J., Kikuchi, N., Kojima, Y., Koya, S., Kusakabe, M., Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, Y., Owa, C., Ozawa, Y., Saito, H., Sano, M., Sato, K., Shibata, K., Shibata, Y., Shigemoto, Y., Shiraki, T., Sogabe, Y., Sugahara, Y., Suzuki, H., Suzuki, H., Takahashi, F., Tateno, M., Tominaga, N., Tsunoda, Y., Wachihi, A., Watanabe, S., Yamamura, T., Yasunishi, A., Yokota, T., Yoshiki, A., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
 RIKEN Mouse ESTs (Konno, H., et al. 1999)
 Unpublished (1999)
 Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@sc.riken.jp, URL: http://genome.gsc.riken.jp/
 Sasaki, N., Izawa, M., Watahiki, M., Ozawa, K., Tanaka, T., Yoneda, Y., Matsuura, S., Carninci, P., Muramatsu, M., Okazaki, Y. and Hayashizaki, Y.
 Transcriptional sequencing: A method for DNA sequencing using RNA polymerase. Proc. Natl. Acad. Sci. U.S.A. 95 (7), 3455-3460 (1998)
 Itoh, M., Kitsuai, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J., Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki, Y. and Hayashizaki, Y.
 Automated filtration-based high-throughput plasmid preparation system. Genome Res. 9 (5), 463-470 (1999)
 Carninci, P. and Hayashizaki, Y.
 High-efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)
 Please visit our web site (http://genome.rtc.riken.go.jp) for further details.

FEATURES
 source Location/Qualifiers
 1..274
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="4930406D09"
 /sex="male"
 /tissue_type="testis"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="RIKEN full-length enriched, adult male testis (DH10B)"
 /notes="Site 1: SalI; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken

contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5'
 GAGAGAGAGATCCAGAGCTCTTTTTTTTTTTTTTTTNN 3'], cDNA was prepared by using thermostable thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5'
 GAGAGAGATCTCGAGTTAATTAATTAATTCCTCCCTCCCTCCCTCC 3']. cDNA was cloned into the XhoI and BamHI sites. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI; 3' end: BamHI."

ORIGIN
 Query Match 93.6%; Score 23.4; DB 1; Length 274;
 Best Local Similarity 96.0%; Pred. No. 5.7;
 Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 AAAAACTGCATCCCAATGTTATG 25
 |||||
 Db 216 AAAAACTGCATGCCCAATGTTATG 240

RESULT 3
 AV32113 276 bp mRNA linear EST 11-NOV-1999
 LOCUS AV32113 RIKEN full-length enriched, adult male medulla oblongata
 DEFINITION Mus musculus cDNA clone 6330530C15 3' similar to X62840 R.rattus mRNA for potassium channel protein (3145 bp), mRNA sequence.

ACCESSION AV32113
 VERSION AV32113.1 GI:6372165
 KEYWORDS Mus musculus (house mouse)
 SOURCE Mus musculus
 ORGANISM Mus musculus

REFERENCE
 AUTHORS Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 276)
 Konno, H., Aizawa, K., Akahira, S., Akiyama, J., Carninci, P., Endo, T., Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N., Hirozane, T., Hori, F., Ishii, Y., Ishikawa, T., Itoh, M., Izawa, M., Kadota, K., Kagawa, I., Kai, C., Kawai, J., Kikuchi, N., Kojima, Y., Koya, S., Kusakabe, M., Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, Y., Owa, C., Ozawa, Y., Saito, H., Sano, M., Sato, K., Shibata, K., Shibata, Y., Shigemoto, Y., Shiraki, T., Sogabe, Y., Sugahara, Y., Suzuki, H., Suzuki, H., Takahashi, F., Tateno, M., Tominaga, N., Tsunoda, Y., Wachihi, A., Watanabe, S., Yamamura, T., Yasunishi, A., Yokota, T., Yoshiki, A., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
 RIKEN Mouse ESTs (Konno, H., et al. 1999)
 Unpublished (1999)
 Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@sc.riken.jp, URL: http://genome.gsc.riken.jp/
 Sasaki, N., Izawa, M., Watahiki, M., Ozawa, K., Tanaka, T., Yoneda, Y., Matsuura, S., Carninci, P., Muramatsu, M., Okazaki, Y. and Hayashizaki, Y.
 Transcriptional sequencing: A method for DNA sequencing using RNA polymerase. Proc. Natl. Acad. Sci. U.S.A. 95 (7), 3455-3460 (1998)
 Itoh, M., Kitsuai, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J., Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki, Y. and Hayashizaki, Y.
 Automated filtration-based high-throughput plasmid preparation system. Genome Res. 9 (5), 463-470 (1999)
 Carninci, P. and Hayashizaki, Y.
 High-efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)
 Please visit our web site (http://genome.rtc.riken.go.jp) for further details.

FEATURES
 source Location/Qualifiers
 1..274
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="4930406D09"
 /sex="male"
 /tissue_type="testis"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="RIKEN full-length enriched, adult male testis (DH10B)"
 /notes="Site 1: SalI; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken

Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata,K., Shingawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S., Rogers,J., Birney,E. and Hayashizaki,Y.
 Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)
 12466851
 Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-resgsc.riken.jp, URL:http://genome.gsc.riken.jp/

TITLE
 JOURNAL
 PUBMED
 COMMENT

Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hirozane,T., Imotani,K., Ishii,Y., Itoh,M., Kawai,J., Konno,H., Miyazaki,A., Murata,M., Nakamura,M., Nomura,K., Numazaki,R., Ohno,M., Sakai,K., Sakazume,N., Sasaki,D., Sato,K., Shibata,K., Shiraki,T., Tagami,M., Waki,K., Watahiki,A., Muramatsu,M. and Hayashizaki,Y. Direct Submission
 Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
 Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
 RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
 Tissues were provided by Michela Fagiolini and Takao K. Hensch (Laboratory for Neuronal Circuit Development Brain Science Institute RIKEN 2-1 Hirosewa,Wako-shi,Saitama 351-0198 Japan) whose assistance we gratefully acknowledge.

Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES
 source

Location/Qualifiers
 1. .413
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="K330041H11"
 /tissue_type="visual cortex"
 /clone_lib="RIKEN full-length enriched, visual cortex"

ORIGIN

Query Match 93.6%; Score 23.4; DB 5; Length 413;
 Best Local Similarity 96.0%; Pred. No. 6.2;
 Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AAAAACTGCATCCCCCAATGTTATG 25
 |||||
 Db 337 AAAAACTGCATGCCCAATGTTATG 361

RESULT 6
 AI847234/c

LOCUS AI847234 419 bp mRNA linear EST 15-JUL-1999
 DEFINITION UI-M-A11-afq-d-04-0-UI.s1 NIH-BMAP_MBS_N Mus musculus cDNA clone
 UI-M-A11-afq-d-04-0-UI 3', mRNA sequence.
 ACCSSION AI847234
 VERSION AI847234.1 GI:5491140
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)

ORGANISM

Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridea; Muridae; Murinae; Mus.
 1 (bases 1 to 419)

REFERENCE
 AUTHORS
 TITLE

Bonaldo,M.F., Lennon,G. and Soares,M.B.
 Normalization and subtraction: two approaches to facilitate gene discovery
 Genome Res. 6 (9), 791-806 (1996)

JOURNAL
 PUBMED
 COMMENT

Contact: Chin, H
 National Institute of Mental Health
 6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
 20892-9643, USA
 Tel: 301 443 1706
 Fax: 301 443 9890

Email: mES@mail.nih.gov
 The sequence contained an oligo-dT track that was present in the oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A tail. The sequence tag present in the cDNA between the NotI site and the oligo-dT track served to verify it as a clone from the normalized brain stems library cDNA library Preparation: M.B. Soares Lab Clone distribution: NIH-BMAP cDNA clones will be made available by the means that is soon to be determined. When NIH determines the means for distribution of the BMAP cDNA clones, this record will be updated accordingly when that means is determined. The following repetitive elements were found in this cDNA sequence: 180-217, >POLY_A\$implete
 Seq primer: M13 Forward
 POLYA=Yes.

FEATURES
 source

Location/Qualifiers
 1. .419
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UI-M-A11-afq-d-04-0-UI"
 /dev_stage="27-32 days"
 /lab_host="PH10B (Life Technologies)"
 /clone_lib="NIH-BMAP_MBS_N"
 /note="vector: pRT3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; The NIH-BMAP_MBS_N library is a normalized library constructed from mouse brain stems. The tag is a string of 5 nucleotides present between the Not I site and the oligo-dT track. The library was constructed as described by Bonaldo, Lennon and Soares, Genome Research 6: 791-806, 1996. Tissue provided by Ms. Annie Novakovich, Zivic-Miller Laboratories.
 TAG_TISSUE=brain-stems
 TAG_LIB=NIH-BMAP_MBS_N
 TAG_SEQ=TCATG"

ORIGIN

Query Match 93.6%; Score 23.4; DB 1; Length 419;
 Best Local Similarity 96.0%; Pred. No. 6.2;
 Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AAAAACTGCATCCCCCAATGTTATG 25
 |||||
 Db 185 AAAAACTGCATGCCCAATGTTATG 161

RESULT 7
 BY616767

LOCUS BY616767 438 bp mRNA linear EST 15-DEC-2002
 DEFINITION RIKEN full-length enriched, visual cortex Mus musculus
 cDNA clone K330021F10 3', mRNA sequence.
 ACCSSION BY616767
 VERSION BY616767.1 GI:26951949
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)

ORGANISM	Mus musculus	source	1. .438
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Mus.		/organism="Mus musculus"
AUTHORS	Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nikaide, I., Oato, N., Saito, R., Saito, R., Suzuki, H., Yamana, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Shonbach, C., Gojbori, T., Baldarelli, R., Hill, D.P., Bult, C., Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H., Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusic, V., Chothia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A., Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W.J., Pertea, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramchandran, S., Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M., Sandelin, A., Schneider, C., Sempile, C.A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S., Rogers, J., Birney, E. and Hayashizaki, Y.		/mol_type="mRNA"
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs		/strain="C57BL/6J"
JOURNAL	Nature 420, 563-573 (2002)		/db_xref="taxon:10090"
PUBLISHED	12466851		/clone="K330021F10"
COMMENT	Contact: Yoshihide Hayashizaki Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute The Institute of Physical and Chemical Research (RIKEN) 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan Tel: 81-45-503-9222 Fax: 81-45-503-9216 Email: genome-res@genome.riken.jp, URL: http://genome.gsc.riken.jp/ Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001) Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000) RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000) Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001) cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Tissues were provided by Michela Fagioli and Takao K. Hensch (Laboratory for Neuronal Circuit Development Brain Science Institute RIKEN 2-1 Hiroo, Wako-shi, Saitama 351-0198 Japan) whose assistance we gratefully acknowledge. Please visit our web site (http://genome.gsc.riken.go.jp) for further details.		/db_xref="taxon:10090"
			/tissue_type="visual cortex"
			/clone_lib="RIKEN full-length enriched, visual cortex"
ORIGIN			
Query Match	93.6%; Score 23.4; DB 5; Length 438;		
Best Local Similarity	96.0%; Pred. No. 6.2;		
Matches	24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;		
QY	1 AAAAATGTCATCCCAATGTTATG 25		
DB	276 AAAAATGTCATCCCAATGTTATG 300		
RESULT 8			
BB750064			
LOCUS	BB750064 RIKEN full-length enriched, pooled tissues, cerebellum, etc. Mus musculus cDNA clone G130001l13 3', mRNA sequence.		
DEFINITION	BB750064		
ACCESSION	BB750064.1 GI:16154300		
VERSION	EST.		
KEYWORDS	Mus musculus (house mouse)		
SOURCE	Mus musculus		
ORGANISM	Mus musculus		
REFERENCE	1 (bases 1 to 449)		
AUTHORS	Okazaki, Y., Arakawa, T., Carninci, P., Furuno, M., Hanagaki, T., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Imotani, K., Ishii, Y., Ito, M., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Matsuyama, T., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Watahiki, A., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.		
TITLE	RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura, T., et al. 2001)		
JOURNAL	Unpublished (2001)		
COMMENT	Contact: Yoshihide Hayashizaki Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute The Institute of Physical and Chemical Research (RIKEN) 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan Tel: 81-45-503-9222 Fax: 81-45-503-9216 Email: genome-res@genome.riken.jp, URL: http://genome.gsc.riken.jp/ Carninci, P., Shibata, Y., Hayata, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000) wagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000) Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara, Y. and Hayashizaki, Y. Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001) Please visit our web site (http://genome.gsc.riken.go.jp) for further details. e mouse tissues.		
FEATURES			

FEATURES
source

Location/Qualifiers
1. .449
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="G13000113"
/clone_lib="RIKEN full-length enriched, pooled tissues, cerebellum, etc."
/notes="pooled tissues ; (tissue_type=cerebellum, dev_stage=16 days neonate, sex=mixed), (tissue_type=cerebellum, dev_stage=0 day neonate, sex=mixed), (tissue_type=hippocampus, dev_stage=adult, sex=male), (tissue_type=whole body, dev_stage=9 days embryo, sex=mixed), (tissue_type=lung, dev_stage=13 days embryo, sex=mixed)"

ORIGIN

Query Match 93.6%; Score 23.4; DB 2; Length 449;
Best Local Similarity 96.0%; Pred. No. 6.3;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAAACTGCATCCCAATGTTATG 25
|||||
Db 57 AAAAACTGCATCCCAATGTTATG 81
|||||

RESULT 9

AI844638/c
LOCUS
DEFINITION
UI-M-ALI-ahr-b-07-0-UI.s1 NIH BMAP MCO N Mus musculus cDNA clone
UI-M-ALI-ahr-b-07-0-UI 3', mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
EST.
AI844638.1 GI:5488544

Mus musculus (house mouse)

ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.

REFERENCE
AUTHORS
TITLE
Normalization and subtraction: two approaches to facilitate gene discovery
Bonaldo M.F., Lennon G. and Soares M.B.
Genome Res. 6 (9), 791-806 (1996)

8889548

JOURNAL
PUBMED
COMMENT
Contact: Chin, H
National Institute of Mental Health
6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
20892-9643, USA
Tel: 301 443 1706
Fax: 301 443 9890

Email: mEST@mail.nih.gov
The sequence contained an oligo-dT track that was present in the oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A tail. The sequence tag present in the cDNA between the NotI site and the oligo-dT track served to verify it as a clone from the normalized prefrontal cortex library cDNA Library Preparation: M.B. Soares Lab Clone distribution: NIH BMAP cDNA clones will be made available by the means that is soon to be determined. When NIH determines the means for distribution of the BMAP cDNA clones, this record will be updated accordingly when that means is determined. The following repetitive elements were found in this cDNA sequence:
169-206, >POLY A#Simple_repeat
Seq primer: M13 Forward
POLYA=yes.

FEATURES

Location/Qualifiers
1. .474
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"

source

/clone="UI-M-ALI-ahr-b-07-0-UI"
/dev_stage="27-32 days"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NIH BMAP MCO_N"
/notes="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site1: NotI; Site 2: Eco RI; The NIH BMAP MCO_N library is a normalized library constructed from mouse cortex. The tag is a string of 5 nucleotides present between the NotI site and the oligo-dT track. The library was constructed as described by Bonaldo, Lennon and Soares, Genome Research 6: 791-806, 1996. Tissue provided by Ms. Annie Novakovich, Zivic-Miller Laboratories.
TAG_TISSUE=prefrontal-cortex
TAG_LIB=NIH BMAP_MCO_N
TAG_SEQ=GTCTCA"

ORIGIN

Query Match 93.6%; Score 23.4; DB 1; Length 474;
Best Local Similarity 96.0%; Pred. No. 6.3;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAAACTGCATCCCAATGTTATG 25
|||||
Db 174 AAAAACTGCATCCCAATGTTATG 150
|||||

RESULT 10

BB130746
LOCUS
DEFINITION
BB130746 RIKEN full-length enriched, 16 days neonate cerebellum Mus musculus cDNA clone 9630047A19 3', mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
EST.
BB130746.2 GI:16266316

Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.

REFERENCE

Arakawa T., Carninci P., Fukuda S., Furuno M., Hangaki T., Hara A., Hiramoto K., Hori F., Ishii Y., Ito M., Kawai J., Konno H., Kouda M., Koya S., Matsuyama T., Miyazaki A., Nomura K., Ohno M., Okazaki Y., Okido T., Saito R., Sakai C., Sakai K., Sano H., Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takanashi F., Takeda Y., Tanaka T., Toyota T., Muramatsu M. and Hayashizaki Y. RIKEN Mouse ESTs (Arakawa T., et al. 2001)
Unpublished (2001)

TITLE

JOURNAL

COMMENT

On Jun 28, 2000 this sequence version replaced gi:8785136.
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-resgsc.riken.jp, URL: http://genome.gsc.riken.jp/
Itoh M., Konno H., Okazaki Y., Muramatsu M. and Hayashizaki Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
wagi K., Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M., Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J., Okazaki Y., Muramatsu M., Inoue Y., Kira A. and Hayashizaki Y.
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)
Konno H., Fukunishi Y., Shibata K., Itoh M., Carninci P., Sugahara Y. and Hayashizaki Y.

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanaka,I., Aizawa,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and Hayashizaki,Y.
 Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
 Please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.
 cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

FEATURES

source

Location/Qualifiers
 1. .648
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="9630047A19"
 /issue_type="cerebellum"
 /dev_stage="16 days neonate"
 /lab_host="DH10B"
 /clone_lib="RIKEN full-length enriched, 16 days neonate cerebellum"
 /note="Site_1: SalI; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGAGAGAGATCCAGACTCTTTTCTTTTCTTTT 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 20.0 and subtraction to Rot = 370.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGATTCGAGTTAAATTAATTCCTCCCCCCCC 3']. cDNA was cleaved with XhoI and BamHI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I."

ORIGIN

Query Match 93.6%; Score 23.4; DB 1; Length 648;
 Best Local Similarity 96.0%; Pred. No. 6.7;
 Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AAAAACTGCATCCCCCAATGTTATG 25

Db 256 AAAAACTGCATGCCCAATGTTATG 280

RESULT 11

AK079373

LOCUS

AK079373 887 bp mRNA linear HTC 03-APR-2004
 Mus musculus 16 days neonate cerebellum cDNA, RIKEN full-length enriched library, clone:9630047A19 product:unknown EST, full insert sequence.

ACCESSION

AK079373

VERSION

AK079373.1 GI:26098461

KEYWORDS

HTC; CAP trapper.

SOURCE

Mus musculus (house mouse)

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

Sciurognathi; Muridae; Murinae; Mus.

1

REFERENCE

AUTHORS

Carninci,P. and Hayashizaki,Y.

TITLE

High-efficiency full-length cDNA cloning

JOURNAL

Meth. Enzymol. 303, 19-44 (1999)

PUBMED

10349636

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

2 Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Kono,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000)
 11042159

3 Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Kono,H., Akiyama,J., Nishi,K., Kitsuina,T., Tashiro,H., Itoh,M., Suni,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
 RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
 Genome Res. 10 (11), 1757-1771 (2000)
 11076861

4 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
 Functional annotation of a full-length mouse cDNA collection
 Nature 409, 685-690 (2001)

5 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
 Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)

6 (bases 1 to 887)
 Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T., Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Kono,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.
 Direct Submission

Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.jp, URL:<http://genome.gsc.riken.jp/>, Tel:81-45-503-9222, Fax:81-45-503-9216)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.

URL:<http://genome.gsc.riken.jp/>URL:<http://fantom.gsc.riken.jp/>

Location/Qualifiers

1. .887

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="FANTOM_DB:9630047A19"

/db_xref="taxon:10090"

/clone="9630047A19"

/issue_type="cerebellum"

/clone_lib="RIKEN full-length enriched mouse cDNA library"

/dev_stage="16 days neonate"

1. .887

misc_feature

/note="unknown EST (GB|BE750064, evidence: BLASTN, 99%, match=451)"

ORIGIN

Query Match 93.6%; Score 23.4; DB 4; Length 887;
 Best Local Similarity 96.0%; Pred. No. 7.2; 1; Indels 0; Gaps 0;
 Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAAACTGCATCCCAATGTTATG 25
 |||||
 Db 495 AAAAACTGCATCCCAATGTTATG 519

RESULT 12

BB083460 175 bp mRNA linear EST 28-JUN-2000
 LOCUS BB083460 RIKEN full-length enriched, adult male diencephalon Mus
 DEFINITION musculus cDNA clone 9330182L24 3', mRNA sequence.

ACCESSION BB083460
 VERSION BB083460
 KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM

Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 (bases 1 to 175)
 Konno,H., Aizawa,K., Akahira,S., Akiyama,J., Arakawa,T., Hayatsu,N.,
 Carninci,P., Endo,T., Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N.,
 Hirozane,T., Hori,F., Ishii,Y., Ishikawa,J., Ishikawa,T., Itoh,M.,
 Izawa,M., Kadota,K., Kagawa,I., Kai,C., Kawai,J., Kikuchi,N.,
 Kiyosawa,H., Kojima,Y., Kondo,S., Koya,S., Kurihara,C.,
 Kusakabe,M., Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H.,
 Okazaki,Y., Ono,T., Owa,C., Saito,H., Sakai,C., Sato,K.,
 Shibata,K., Shibata,Y., Shigemoto,Y., Shinagawa,A., Shiraki,T.,
 Sugabe,Y., Sugahara,Y., Suzuki,H., Suzuki,H., Tagawa,A.,
 Takahashi,F., Tomimaga,N., Toya,T., Tsunoda,Y., Watahiki,A.,
 Watanabe,S., Yamamura,T., Yamanaka,I., Yano,R., Yasunishi,A.,
 Yokota,T., Yoshida,K., Yoshihiki,A., Yoshino,M., Muramatsu,M. and
 Hayashizaki,Y.
 RIKEN Mouse ESTs (Konno,H., et al.)
 Unpublished (2000)

TITLE

JOURNAL

COMMENT

Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@sc.riken.jp, URL: http://genome.gsc.riken.jp/
 Carninci,P., Nishiyama,Y., Westover,A., Itoh,M., Nagaoka,S.,
 Sasaki,N., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
 Thermotabilization and thermoactivation of thermolabile enzymes by
 trehalose and its application for the synthesis of full length
 cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
 Itoh,M., Kitsumai,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J.,
 Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M.,
 Okazaki,Y. and Hayashizaki,Y.
 Automated filtration-based high-throughput plasmid preparation
 system. Genome Res. 9 (5), 463-470 (1999)
 Carninci,P. and Hayashizaki,Y.
 High-efficiency full-length cDNA cloning. Methods Enzymol. 303,
 19-44 (1999)
 Please visit our web site (http://genome.rtc.riken.go.jp) for
 further details.

FEATURES

source

Location/Qualifiers

1..175
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="C57BL/6J"
 /clone="9330182L24"
 /sex="male"
 /tissue_type="diencephalon"
 /dev_stages="adult"

ORIGIN

Query Match 89.6%; Score 22.4; DB 1; Length 220;
 Best Local Similarity 95.8%; Pred. No. 16;
 Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

/lab_host="DH10B"
 /clone_lib="RIKEN full-length enriched, adult male
 diencephalon"
 /note="Site 1: Sali; Site 2: BamHI; cDNA library was
 prepared and sequenced in Mouse Genome Encyclopedia
 Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in
 RIKEN. Division of Experimental Animal Research in Riken
 contributed to prepare mouse tissues. 1st strand cDNA was
 primed with a primer [5'
 GAGAGAGAGAGATCCAGAGCTCTTTTTTTTTTTTNN 3'], cDNA was
 prepared by using trehalose thermo-activated reverse
 transcriptase and subsequently enriched for full-length by
 cap-trapper. cDNA went through one round of normalization
 to Rot = 10.0 and subtraction to Rot = 185.0. Second
 strand cDNA was prepared with the primer adapter of
 sequence [5' GAGAGAGATTCGAGTTAAATTAATCCCCCCCCCCC
 3']. cDNA was cloned into the XhoI and BamHI sites.
 Vector: a modified pBluescript KS(+) after bulk excision
 from Lambda FLC I. Cloning sites, 5' end: Sali; 3' end:
 BamHI"

ORIGIN

Query Match 89.6%; Score 22.4; DB 1; Length 175;
 Best Local Similarity 95.8%; Pred. No. 15;
 Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAAACTGCATCCCAATGTTAT 24

Db 113 AAAAACTGCATCCCAATGTTAT 136

RESULT 13

AI412710/C

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

Location/Qualifiers

1..220

/organism="Rattus sp."

/mol_type="mRNA"

/db_xref="taxon:10118"

/clone="RBRDS47"

/notes="Organ: brain; Vector: pT73Pac; Site 1: EcoRI;
 Site 2: NotI"

Seq primer: M13-21.

Contact: Lee, NH

The Institute for Genomic Research

9712, Medical Center Drive, Rockville, MD 20850, USA

Tel: (301)-838-3529

Fax: (301)-838-0208

Email: nhlee@tigr.org

Unpublished (1998)

Rat Genome Project: Generation of a Rat EST (REST) Catalog & Rat

Gene Index

Lee,N.H., Glodok,A., Chandra,I., Mason,T.M., Quackenbush,J.,

Kerlavage,A.R. and Adams,M.D.

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

Sciurognathi; Muridae; Murinae; Rattus.

1 (bases 1 to 220)

AI412710.1 GI:4256214

Rattus sp.

Rattus sp.

EST.

EST241009 Normalized rat brain, Bento Soares Rattus sp. cDNA clone

RBRDS47 3' end, mRNA sequence.

AI412710

Genetics This clone is also available through the I.M.A.G.E. Consortium at LNL (inf@image.llnl.gov). IMAGE ID=1773693 The following repetitive elements were found in this cDNA sequence: 75-114, >POLY_A#Simple_repeat
Seq primer: M13 Forward
POLYA=No.

FEATURES
source
1. .380
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-CQ-hq-f-10-0-UI"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="UI-R-CO"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; The UI-R-CO library is a subtracted library derived from the UI-R-A1 and UI-R-E1 libraries. The UI-R-A1 library consisted of a mixture of individually tagged normalized libraries constructed from rat placenta, adult lung, brain, liver, kidney, heart, spleen, ovary, and muscle. The UI-R-E1 library consisted of a mixture of individually tagged normalized libraries constructed from 8, 12 and 18-day embryo. The tag is a string of 3-5 nucleotides present between the Not I site and the oligo-dT track which allows identification of the library of origin of a clone within the mixture. The subtracted library (UI-R-CO) was constructed as follows: PCR amplified cDNA inserts from a pool of UI-R-A1 and UI-R-E1 clones from which 3' ESTs had been derived was used as a driver in a hybridization with the pooled UI-R-A1 and UI-R-E1 library in the form of single-stranded circles. The remaining single-stranded circles (subtracted library) was purified by hydroxyapatite column chromatography, converted to double-stranded circles and electroporated into DH10B bacteria (Life Technologies) to generate the UI-R-CO library. This procedure has been previously described (Bonaldo, Lennon and Soares, Genome Research 6: 791-806, 1996)"

ORIGIN
Query Match 89.6%; Score 22.4; DB 1; Length 380;
Best Local Similarity 95.8%; Pred. No. 18;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 AAAAACTGCATCCCAATGTTATG 25
Db 79 AAAAACTGCATCCCAATGTTATG 56
|||||

RESULT 18
BQ177061/c
LOCUS
DEFINITION
UI-M-DJ2-bwa-b-13-0-UI.s1 NIH_BMAP_DJ2 Mus musculus cDNA clone
BQ177061
ACCESSION
VERSION
KEYWORDS
SOURCE
Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 512)
Bonaldo,M.F., Lennon,G. and Soares,M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
8889548
PUBMED
COMMENT
Contact: Chin, H
National Institute of Mental Health

6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
20892-9643, USA
Tel: 301 443 1706
Fax: 301 443 9890
Email: mEST@mail.nih.gov
Tissue Procurement: Dr. Robin Davisson
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).
The following repetitive elements were found in this cDNA
sequence: 170-207, >POLY_A#Simple_repeat (matched compliment)
Seq primer: M13 FORWARD
POLYA=Yes.

FEATURES
source
1. .512
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="UI-M-DJ2-bwa-b-13-0-UI"
/tissue="subfornical organ and postrema"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="NIH_BMAP_DJ2"
/note="Organ: brain; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; UI-M-DJ2 is a subtracted cDNA library constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is: GCTACATGAT, subfornical organ and area postrema.
TAG_TISSUE=subfornical organ and postrema
TAG_LIB=UI-M-DJ2
TAG_SEQ=GCTACATGAT"

ORIGIN
Query Match 87.2%; Score 21.8; DB 3; Length 512;
Best Local Similarity 92.0%; Pred. No. 36;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 AAAAACTGCATCCCAATGTTATG 25
Db 175 AAAAACTGCATCCCAATGTTTG 151
|||||

RESULT 19
BB629036
LOCUS
DEFINITION
BB629036 RIKEN full-length enriched, 16 days neonate cerebellum Mus
musculus cDNA clone 9630047A19 5', mRNA sequence.
BB629036
ACCESSION
VERSION
KEYWORDS
SOURCE
Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 647)
Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T.,
Harai,A., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J.,
Konno,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K.,
Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K.,
Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,

Takeda,Y., Tanaka,T., Toya,T., Muramatsu,M. and Hayashizaki,Y.
RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
Unpublished (2001)
Contact: Yoshinhide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Teurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-resesc.riken.jp, URL:http://genome.gsc.riken.jp/
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
wagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T.,
Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A.
and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P.,
Sugahara,Y. and Hayashizaki,Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanaka,I.,
Aizawa,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and
Hayashizaki,Y.
Computational Analysis of Full-Length Mouse cDNAs Compared with
Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)
Please visit our web site (http://genome.gsc.riken.go.jp) for
further details.
e mouse tissues.
FEATURES
source
1. .647
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="9630047A19"
/tissue_type="cerebellum"
/dev_stage="16 days neonate"
/lab_host="DRI0B"
/clone_lib="RIKEN full-length enriched, 16 days neonate
cerebellum"
/notes="Site 1: Sali; Site 2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5'
GAGAGAGAGAGATCCAGAGCTCTTTTCTTTTCTTNN 3'], cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. cDNA went through one round of normalization
to Rot = 20.0 and subtraction to Rot = 370.0. Second
strand cDNA was prepared with the primer adapter of
sequence [5' GAGAGAGAGATCTCGAGTTAATAATTAATCCCCCCCCC
3']. cDNA was cleaved with XhoI and BamHI. Vector: a
modified pBluescript KS(+) after bulk excision from Lambda
FLC I."

RESULT 20
COB73358/c
LOCUS
DEFINITION
BovGen_01683 normal cattle brain Bos taurus CDNA clone
RZPDP1056N1259Q 3', mRNA sequence.
ACCESSION
COB73358
VERSION
COB73358.1 GI:51803198
KEYWORDS
EST.
SOURCE
Bos taurus (cow)
ORGANISM
Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
Pecora; Bovidae; Bovinae; Bos.
1 (bases 1 to 744)
Hennig,S., Janitz,M., Herwig,R. and Williams,J.
Generation, annotation, evolutionary analysis and database
integration of 14969 cattle EST clusters
Unpublished (2004)
JOURNAL
Contact: Hennig S
Laboraty 123, dept. Lehrach
Max-Planck-Institut fuer Molekulare Genetik
Inhnstr.63-73, D-14195 Berlin, Germany
Tel: +49 30 8413 1612
Fax: +49 30 8413 1380
Email: hennig@molgen.mpg.de
The library was characterised by oligonucleotide fingerprinting
(ONFP) to reduce sequencing redundancy. According to the ONFP
procedure, clones that display the same hybridisation matrix with a
battery of 200 8mer oligonucleotides are grouped into clusters. One
clone per ONFP cluster was selected for sequencing. cDNA clones and
filters are distributed via Deutsches Ressourcenzentrum fuer
Genomforschung GmbH (http://www.rzpd.de).
PCR Primers
FORWARD: 5' CCCAGGCTTTACATTTATGTTCCGGCTCG 3' (M13RSP) 5'-seq
BACKWARD: 5' GCTATTACGCCAGCTGGGAAAGGGGATGTG 3' (M13FSP) 3'-seq
Seq primer: 5' GCTATTACGCCAGCTGGGAAAGGGGATGTG 3' (M13FSP).
FEATURES
source
1. .744
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/clone="RZPDP1056N1259Q"
/sex="female"
/tissue_type="brain tissue"
/dev_stage="adult brain"
/clone_lib="normal cattle brain"
/note="Organ: brain; Vector: pSport1; Site 1: NotI;
Site 2: Sali; Random primed and directionally cloned in
pSport1 vector using NotI
(5'-pCAGTACTTCTAGATCGGAGCGCGGCC (T)15-3' and Sali 5'-
TGCACCCACGCGGTCCG-3' adapters (Gibco BRL)"]
ORIGIN
Query Match 87.2%; Score 21.8; DB 7; Length 744;
Best Local Similarity 92.0%; Pred. No. 39;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 AAAAACTGCATCCCAATGTTATG 25
||||| ||||| ||||| ||||| |||||
DB 413 AAAAACTGCATCCCAATGTTATG 389
||||| ||||| ||||| ||||| |||||
RESULT 21
COB78266
LOCUS
DEFINITION
BovGen_06591 normal cattle brain Bos taurus CDNA clone
RZPDP1056K1223Q 5', mRNA sequence.
ACCESSION
COB78266
VERSION
COB78266.1 GI:51808182
KEYWORDS
EST.
SOURCE
Bos taurus (cow)

Mon Feb 6 12:23:16 2006

```

ORGANISM      Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
Pecora; Bovidae; Bovinae; Bos.
REFERENCE
AUTHORS      Hennig,S., Janitz,M., Herwig,R. and Williams,J.
TITLE        Generation, annotation, evolutionary analysis and database
              integration of 14969 cattle EST clusters
JOURNAL
COMMENT      Unpublished (2004)
              Contact: Hennig S
              Laboratory 123, dept.Lehrach
              Max-Planck-Institut fuer Molekulare Genetik
              Ihnestr.63-73, D-14195 Berlin, Germany
              Tel: +49 30 8413 1612
              Fax: +49 30 8413 1380
              Email: hennig@molgen.mpg.de
              The library was characterised by oligonucleotide fingerprinting
              (ONFP) to reduce sequencing redundancy. According to the ONFP
              procedure, clones that display the same hybridisation matrix with a
              battery of 200 8mer oligonucleotides are grouped into clusters. One
              clone per ONFP cluster was selected for sequencing. cDNA clones and
              filters are distributed via Deutsches Ressourcenzentrum fuer
              Genomforschung GmbH (http://www.rzpd.de).
              PCR Primers
              FORWARD: 5' CCCAGGCTTTACACTTTATGCTCGGCTCG 3' (M13RSP) 5'-seq
              BACKWARD: 5' SCTATTACGACGCTGGCGAAGGGGATGTG 3' (M13FSP) 3'-seq
              Seq primer: 5'-CCGTCGCGAATTCGGGT-3' (M13RSP).
              Location/Qualifiers
                1. .764
                  /organism="Bos taurus"
                  /mol_type="mRNA"
                  /db_xref="taxon:9913"
                  /clone="RZPDp1056K1223Q"
                  /sex="female"
                  /tissue type="brain tissue"
                  /dev stage="adult brain"
                  /clone lib="normal cattle brain"
                  /note="Organ: brain; Vector: pSport1; Site 1: NotI;
                  Site 2: SalI; Random primed and directionally cloned in
                  pSport1 vector using NotI
                  (5'-pGACTAGTTCTAGATCGCGCGCGGCC (T)15-3' and SalI 5'-
                  TCGACCCACGCGTCCG-3' adapters (Gibco BRL))"

ORIGIN
Query Match      87.2%; Score 21.8; DB 7; Length 764;
Best Local Similarity 92.0%; Pred. No. 39;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAACTGCATCCCAATGTTATG 25
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 520 AAAAACTGCATGCCCAATGTTATG 544

RESULT 22
BG727345
LOCUS           407 bp mRNA linear EST 09-MAY-2001
DEFINITION     fo80f07.x1 zebrafish gridded kidney Danio rerio cDNA clone
IMAGE:4726428 3', mRNA sequence.
ACCESSION      BG727345
VERSION        BG727345.1 GI:14012420
KEYWORDS       EST.
SOURCE         Danio rerio (zebrafish)
ORGANISM       Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
REFERENCE
AUTHORS        Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Marra,M.,
Eddy,S., Hallier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y.,
Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R.,
Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
Waterston,R. and Wilson,R.

TITLE
JOURNAL
COMMENT      Unpublished (1998)
              Contact: Stephen L. Johnson
              Washington University School of Medicine
              4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
              Tel: 314 286 1800
              Fax: 314 286 1810
              Email: zbrafish@watson.wustl.edu
              cDNA Library Preparation: Leonard Ira Zon DNA Sequencing by:
              Washington University Genome Sequencing Center Clone distribution:
              Genome Systems, St. Louis, Missouri (web address:
              www.genomesystems.com) (email contact: info@genomesystems.com) and
              Research Genetics, Huntsville, Alabama (web address:
              www.resgen.com) (email contact: info@resgen.com) and
              RessourcenzentrumPrimarDatenbank, Berlin, Germany (web address:
              www.rzpd.de)
              Seq primer: T7 from Gibco
              High quality sequence stop: 402.
              Location/Qualifiers
                1. 407
                  /organism="Danio rerio"
                  /mol_type="mRNA"
                  /db_xref="taxon:7955"
                  /clone="IMAGE:4726428"
                  /sex="mixed"
                  /tissue type="kidney pooled from 300 wild type adults"
                  /lab_host="XL0LR"
                  /clone_lib="zebrafish gridded kidney"
                  /note="Organ: kidney; Vector: pBK-CMV; Site_1: EcoRI;
                  Site_2: XhoI; Oligo dn cDNA library constructed from mRNA
                  pooled from pooled kidney tissue from 300 adult
                  zebrafish."

ORIGIN
Query Match      84.0%; Score 21; DB 2; Length 407;
Best Local Similarity 100.0%; Pred. No. 83;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AAAAACTGCATCCCAATGTT 22
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 198 AAAAACTGCATCCCAATGTT 218

RESULT 23
BU551162/c
LOCUS           602 bp mRNA linear EST 16-SEP-2002
DEFINITION     GM880019B20D08 Gm-r1088 Glycine max cDNA clone Gm-r1088-7096 3',
mRNA sequence.
ACCESSION      BU551162
VERSION        BU551162.1 GI:22934023
KEYWORDS       EST.
SOURCE         Glycine max (soybean)
ORGANISM       Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
REFERENCE
AUTHORS        Vodkin,L., Shoemaker,R., Keim,P., Retzel,E., Khanna,A., Shealy,R.,
Clough,S., Thibaud-Nissen,F., Coryell,V., Erpeliding,J., Raph,C.,
Shoop,E., Stromvik,M., Schweitzer,P., Gong,G. and Liu,L.
              A Functional Genomics Program for Soybean (NSF 9872565) (2002)
              Unpublished (2002)
              Other_ESTs: BG726542 corresponding to Gm-cl067-4033 (5')
              Contact: Vodkin, L.O., PI, A Functional Genomics Program for
              Soybean (NSF 9872565)
              Lewin, H. A., Director, Keck Center for Comparative and Functional
              Genomics
              University of Illinois
              Edwin R. Madigan Building, 1201 W. Gregory, Urbana, IL 61801, USA
              Tel: (217) 244-6147
              Fax: (217) 333-4582
              Email: l-vodkin@uiuc.edu

```

Insert Length: 602 Std Error: 0.00
 Plate: G880019B20 row: D column: 08
 Seq primer: 5'-TTTTTTTTTTTTTTTTT(A/C/G)-3'
 High quality sequence stop: 602.

FEATURES

source

Location/Qualifiers
 1..602
 /organism="Glycine max"
 /mol_type="mRNA"
 /db_xref="taxon:3847"
 /clone_lib="Gm-r1088-7096"
 /clone_lib="Gm-r1088"
 /note="The library Gm-r1088 is a sequence-driven, reracked set of 9,216 low redundancy clones selected from cDNA libraries from various tissues and stages of development of soybean that consists of 2,706 cDNAs from germinating cotyledons (source library Gm-cl027); 1,355 cDNAs from immature seed coats (libraries Gm-cl019 and Gm-cl023); 917 cDNAs from tissue culture derived somatic embryos (source libraries Gm-cl036 and Gm-cl075); 3,275 cDNAs from germinating seedlings, shoot tips, or leaves exposed to various stresses (source libraries Gm-cl064, Gm-cl065, Gm-cl066, and Gm-cl067); and 963 cDNAs from young leaves exposed to bacterial and fungal pathogens (source libraries Gm-cl072, Gm-cl073, and Gm-cl074). The 5' ESTs of the source clones from the different libraries was used to select singletons, or a representative of each contig, which were reracked to form library Gm-r1088 and the cDNA clones of the reracked Gm-r1088 library were then sequenced at the 3' end. The unigene selection and 3' sequencing was funded by NSF Plant Genome project #9872565 (http://soybeanomics.crops.ci.uiuc.edu/) as part of creation of a low redundancy soybean cDNA set. The source cDNA libraries were constructed by the laboratories of Lila Vodkin, University of Illinois, Randy Shoemaker, Iowa State University, and Paul Keim, Northern Arizona University as part of the Public EST project, http://129.186.26.94/soybeanest.html. The contig analysis to select unique genes was performed by the laboratory of Ernest Retzel, Center for Computational Genomics and Bioinformatics, University of Minnesota, http://web.ahc.umn.edu/biodata/nfssoy/. Reracking and 3' sequencing were conducted by services of the University of Illinois Keck Center for Comparative and Functional Genomics http://www.biotech.uiuc.edu/keck.htm. Note: The library is listed in the 'OTHER EST' field. The detailed information on the source library for each clone can also be obtained by referring to the Genome Systems clone ID of the original cDNA library that is also listed under 'OTHER EST'."

ORIGIN

Query Match 80.8%; Score 20.2; DB 5; Length 602;
 Best Local Similarity 88.0%; Pred. No. 2.1e+02;
 Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAACTGCATCCCAATGTTATG 25

Db 499 ATAAAAATGCATCCCAATGTTAAG 475

RESULT 24

CG818112

LOCUS

DEFINITION SOYEK70TV LargeInsertSoybeanGenLib Glycine max genomic clone

698 bp DNA linear GSS 18-NOV-2003

H65E06:MTP17L19, genomic survey sequence.

ACCESSION

CG818112

VERSION

GSS.

KEYWORDS

SOURCE

ORGANISM

Glycine max (soybean)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
 1 (bases 1 to 698)
 Shultz, J., Meksem, K., Shetty, J., Town, C. D., Koo, H., Potter, J., Wakefield, K., Zhang, H., Wu, C. and Lightfoot, D. A.
 End sequencing of BACs comprising a provisional minimal tiling path from a fingerprint physical map of soybean (Glycine max) cultivar Forrest
 Unpublished (2003)
 Other GSSs: SOYEK70TH
 ContAc: Chris Town, J. L. Shultz and D. A. Lightfoot
 The Center of Excellence in Soybean Research, Teaching and Outreach, Southern Illinois University at Carbondale and Plant Genomics, The Institute for Genomic Research
 Room 176, Ag. Building, Mail Code 4415, Carbondale, IL 62901-4415, USA and 9712 Medical Center Drive, Rockville, MD 20850, USA
 Tel: 618 453 1797
 Email: ga4082@siu.edu, jshultz@siu.edu, cdtown@tigr.org (URL: http://bioinformatics.siu.edu)
 Clones approximating a minimum tiling path were re-arrayed from the library master plates prior to sequencing.
 For purposes of clone identification each clone name is a concatenation of the original clone location and its new location in the re-arrayed sequencing plates.
 Seq primer: GTAATACGACTCACTATAGGGC
 Class: BAC ends.

FEATURES

source

Location/Qualifiers

1..698

/organism="Glycine max"

/mol_type="genomic DNA"

/cultivar="Forrest"

/db_xref="taxon:3847"

/clone_lib="H65E06:MTP17L19"

/note="Organ: Leaves; Vector: pCLD04541 (pBEL0BAC11 EcorI clones); Site: BctXI; Soybean (Glycine max (L.) Merr.) cv. Forrest seeds were grown in greenhouse for fourteen days. Nuclei were isolated and embedded in agarose, restriction digested with Hind III BamHI or EcoRI, large size DNA fragments were ligated in vector V41 (pCLD04541) and electro transformed in DH10a cells. About 90,000 clones from BAC libraries were fingerprinted with HindIII and Hae III. Version 2 (automatic build) Contigs were built from 78,001 fingerprints. Contigs were manually examined to find the best non redundant tile path through the contigs representing about 13,000 clones. The clones were end sequenced."

ORIGIN

Query Match 80.8%; Score 20.2; DB 10; Length 698;
 Best Local Similarity 88.0%; Pred. No. 2.2e+02;
 Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAACTGCATCCCAATGTTATG 25

Db 394 ATAAAAATGCATCCCAATGTTAAG 418

RESULT 25

CO876233

LOCUS

DEFINITION

BovGen 04558 normal cattle brain Bos taurus cDNA clone

R2PDp1056N1259Q 5', mRNA sequence.

ACCESSION

CO876233

VERSION

EST.

KEYWORDS

SOURCE

ORGANISM

Bos taurus (cow)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;

Pecora; Bovidae; Bovinae; Bos.

1 (bases 1 to 726)

Hennig, S., Janitz, M., Herwig, R. and Williams, J.

```

http://image.llnl.gov
plate: LLCM3118 row: i column: 05
High quality sequence stop: 541.
Location/Qualifiers
1. .830
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone="IMAGE:6878382"
/dev_stage="metamorphosis stage 53"
/clone_lib="NICHD_XGC_Radi"
/note="Organ: Developing Tadpole; Vector: pDNR-LIB;
Site 1: Sfi; Site 2: Sfi; 5' and 3' adaptors were used in
cloning as follows: 5' adaptor sequence:
5'-CACGCGCATATGGCC-3' and 3' adaptor sequence:
5'-ATTCTAGAGCCGAGCGCCGACATG-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.6 kb
(range 0.9-3.0 kb). 15/15 colonies contained inserts by
PCR. This library was enriched for full-length clones and
was constructed by Clontech Laboratories (Palo Alto, CA)."
ORIGIN
Query Match 80.8%; Score 20.2; DB 6; Length 830;
Best Local Similarity 88.0%; Pred. No. 2.3e+02;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAACTGCATCCCAATGTTATG 25
|||||
Db 179 AAAAACTGCATACAGATGTTATG 203

RESULT 27
BE555733
LOCUS
DEFINITION BE555733 454 bp mRNA linear EST 13-JUL-2004
Gm-cl045-1181 5', mRNA sequence.
ACCESSION BE555733 GI:9820223
VERSION
KEYWORDS
SOURCE
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eustosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
REFERENCE
1 (bases 1 to 454)
Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V.,
Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J.,
Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M.,
Bowers,X., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N.,
Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,X., Cardenas,M.,
McCann,R., Waterston,R. and Wilson,R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
When it has been determined, an EST from the other end of this
clone is listed in the 'Other ESTs on clone' field. This clone is
available through: Biogenetic Services, 801 32nd Ave. Brookings, SD
57006 USA (phone: 800 423 4163; email: info@biogeneticservices.com)
Insert length: 645 Std Error: 0.00
High quality sequence stop: 411.
Location/Qualifiers
1. .454
/organisms="Glycine max"
/mol_type="mRNA"
/cultivar="Williams 82"
/db_xref="taxon:3847"
FEATURES
source

```


plants of Williams); and 3055 sequences from library Gm-cl028 (from 'Supernod' plants whose seedlings were inoculated with Bradyrhizobium japonicum, courtesy of Dr. Gary Stacey). The 5' ESTs of the source clones from the different progenitor libraries was used to select singletons, or a representative of each contig, which were reracked to form library Gm-r1083. The cDNA clones of the reracked Gm-r1083 library were then sequenced at the 3' end. The contig analysis to select unique genes was performed by the laboratory of Ernest Retzel, Center for Computational Genomics and Bioinformatics, University of Minnesota, http://web.ahc.umn.edu/biodata/nsfsoy/. Reracking was performed by Incyte Genomics, St. Louis, http://www.incyte.com, and 3' sequencing by the Keck Center for Comparative and Functional Genomics, University of Illinois, http://www.lie.uiuc.edu/biotech/keck.html. Note: The corresponding 5' EST from each clone in the Gm-r1083 library is listed in the 'OTHER EST' field. The detailed information on the source library for each clone can also be obtained by referring to the Incyte Genomics clone ID of the original cDNA library that is also listed under 'OTHER EST'."

ORIGIN

Query Match 79.2%; Score 19.8; DB 3; Length 780;
Best Local Similarity 91.3%; Pred. No. 3.5e+02;
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAAAGTCATCCCAATGTTA 23
||||| |||||||
Db 513 ATAAAAATGCATCCCAATGTTA 491

RESULT 30
B1968779/c
LOCUS B1968779.1 G1:16343184
DEFINITION B1968779.1 G1:16343184
mRNA sequence.
B1968779
B1968779.1 G1:16343184
EST.
KEYWORDS
SOURCE Glycine max (soybean)
ORGANISM Glycine max

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
1 (bases 1 to 783)
Vodkin,L., Keim,P., Shoemaker,R., Retzel,E., Khanna,A., Coryell,V., Erpelting,J., Raph,C., Shoop,E., Fardinas,J., Liu,L. and Lewin,H.
A Functional Genomics Program for Soybean (NSF 9872565)
Unpublished (1999)
Other ESTs: AW234285 corresponding to Gm-cl028-804 (5')
Contact: Vodkin, L.O., PI, A Functional Genomics Program for Soybean (NSF 9872565)
Lewin, H. A., Director, Keck Center for Comparative and Functional Genomics
University of Illinois
Edwin R. Madigan Building, 1201 W. Gregory, Urbana, IL 61801, USA
Tel: (217) 244-6147
Fax: (217) 333-4582
Email: l-vodkin@uiuc.edu
This clone is available through: Incyte Genomics, 4633 World Parkway Circle St. Louis, Missouri 63134. Phone (800) 430-0030 or (314) 427-3222 FAX: (314) 427-3324. Web site: http://www.incyte.com/reagents/catalog.jsp?page=clones/collaboration/index

Seq primer: 5'-TTTTTTTTTTTTTTTTT(A/C/G)-3'.
Location/Qualifiers
1. .783
/organism="Glycine max"
/mol_type="mRNA"
/db_xref="taxon:3847"

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

FEATURES

Source
1. .783
/organism="Glycine max"
/mol_type="mRNA"
/db_xref="taxon:3847"

/clone="Gm-r1083-2134"
/clone lib="Gm-r1083"
/note="The library Gm-r1083 is a sequence-driven, reracked set of 4,992 clones selected from cDNA libraries from various tissues and stages of development of soybean. It represents 1117 sequences from the progenitor library Gm-cl009 (from mature roots of 2 month old greenhouse grown 'Williams' soybean plants); 820 sequences from the progenitor library Gm-cl013 (from 2 to 3 week old whole plants of Williams); and 3055 sequences from library Gm-cl028 (from 'Supernod' plants whose seedlings were inoculated with Bradyrhizobium japonicum, courtesy of Dr. Gary Stacey). The 5' ESTs of the source clones from the different progenitor libraries was used to select singletons, or a representative of each contig, which were reracked to form library Gm-r1083. The cDNA clones of the reracked Gm-r1083 library were then sequenced at the 3' end. The contig analysis to select unique genes was performed by the laboratory of Ernest Retzel, Center for Computational Genomics and Bioinformatics, University of Minnesota, http://web.ahc.umn.edu/biodata/nsfsoy/. Reracking was performed by Incyte Genomics, St. Louis, http://www.incyte.com, and 3' sequencing by the Keck Center for Comparative and Functional Genomics, University of Illinois, http://www.lie.uiuc.edu/biotech/keck.html. Note: The corresponding 5' EST from each clone in the Gm-r1083 library is listed in the 'OTHER EST' field. The detailed information on the source library for each clone can also be obtained by referring to the Incyte Genomics clone ID of the original cDNA library that is also listed under 'OTHER EST'."

ORIGIN

Query Match 79.2%; Score 19.8; DB 3; Length 783;
Best Local Similarity 91.3%; Pred. No. 3.5e+02;
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAAAGTCATCCCAATGTTA 23
||||| |||||||
Db 498 ATAAAAATGCATCCCAATGTTA 476

RESULT 31
CK028853
LOCUS CK028853.1 G1:38554777
DEFINITION CK028853.1 G1:38554777
5', mRNA sequence.
CK028853
CK028853.1 G1:38554777
EST.
KEYWORDS
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio

CK028853 407 bp mRNA linear EST 26-NOV-2003
AGENCOURT_16618827 NIH_ZGC_7 Danio rerio cDNA clone IMAGE:7055972
5', mRNA sequence.
CK028853.1 G1:38554777
EST.
KEYWORDS
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Len Zon, Harvard
cDNA Library Preparation: Open Biosystems
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM14835 row: 1 column: 18
High quality sequence stop: 407.

analysis were selected following oligonucleotide hybridization fingerprinting of arrayed clones from zebrafish late somitogenesis (26 ss), adult liver or embryonic shield stage (5, 6 h) libraries. Fingerprint data were used to computationally cluster cDNAs, and a single cDNA from each cluster was chosen for sequencing. In some cases multiple members of the same cluster were sequenced to assess clustering parameters or single clones were sequenced additional times to assess quality control."

ORIGIN

Query Match 77.6%; Score 19.4; DB 1; Length 435;
Best Local Similarity 95.2%; Pred. NO. 4.8e+02;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 AAAAAGTCATCCCCCAATGTT 22
| | | | |
Dd 228 AGAAACTGCATCCCCCAATGTT 248

RESULT 34	AL730907	LOCUS	AL730907	494 bp	mRNA	linear	EST 18-APR-2002
DEFINITION	AL730907 Danio rerio embryonic inner ear subtracted cDNA Danio rerio cDNA clone BN0AA115ZE05 3', mRNA sequence.						

ACCESSION	ALU730507	GI:20195511	
VERSION	ALU730507.1		
KEYWORDS	EST.		
SOURCE	Danio rerio (zebrafish)		
ORGANISM	Danio rerio		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.		

FEATURES
source
1. .494
Location/Qualifiers
Email: sequer@genoscope.cns.fr; web : www.genoscope.cns.fr

ORIGIN

Query Match	77.6%	Score 19.4;	DB 1;	Length 494;
Best Local Similarity	95.2%;	Pred. NO. 4.9e+02;		
Matches 20:	Conservative	0;	Mismatches 1;	Indels 0;
				Gaps 0;

Qy 2 AAAAACTGCATCCCCAATGTT 22
 91 AGAACTGCATCCCCAATGTT 111
 Db

RESULT 35	AL720246/c	AL720246	582 bp	linear	EST 18-APR-2002
LOCUS	AL720246	Danio rerio embryonic inner ear	582 bp	linear	EST 18-APR-2002
DEFINITION	AL720246	embryonic inner ear	582 bp	linear	EST 18-APR-2002
ACCESSION	AL720246	embryonic inner ear	582 bp	linear	EST 18-APR-2002
VERSION	AL720246.1	embryonic inner ear	582 bp	linear	EST 18-APR-2002

KEYWORDS EST.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
REFERENCE 1 (bases 1 to 582)
AUTHORS Coimbra,R., Weil,D., Brottier,P., Blanchard,S., Levi,M.,
Hardelin,J.P., Weissenbach,J. and Petit,C.
TITLE A subtracted cDNA library from the zebrafish (Danio rerio)
embryonic inner ear
JOURNAL Unpublished (2002)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremlieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
Location/Qualifiers
FEATURES 1..582
SOURCE

FEATURES	SOURCE
1. Feature 1	Source 1
2. Feature 2	Source 2
3. Feature 3	Source 3
4. Feature 4	Source 4
5. Feature 5	Source 5
6. Feature 6	Source 6
7. Feature 7	Source 7
8. Feature 8	Source 8
9. Feature 9	Source 9
10. Feature 10	Source 10

ORIGIN

```
Query Match          77.6%; Score 19.4; DB 1; Length 582;
Best Local Similarity 95.2%; Pred. No. 5e+02;
Matches 20: Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

Qy 2 AAAAAGTGCATCCCCAATGTT 22
db 410 AGAAAGTGCATCCCCAATGTT 390

RESULT 36					
AW116002					
LOCUS	AW116002	595 bp	mRNA	linear	EST 20-OCT-1999
DEFINITION	f106a09.x1 Sugano Kawakami zebrafish DRB Danio rerio cDNA clone IMAGE:2600344 3', mRNA sequence.				

VERSION	2W116002.1	GI:6082340
KEYWORDS	EST.	
SOURCE	Danio rerio (zebrafish)	
ORGANISM	Danio rerio	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.	

FEATURES source

Email: hennig@molgen.mpg.de
5' EST sequencing of clones from a zebrafish shield stage library,
normalised from 55,000 starting clones by oligonucleotide
fingerprinting
High quality sequence stop: 802.

FEATURES

Location/Qualifiers
1..802
/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/clone="MPMGp637_16N2;MPMGp637N0216"
/tissue type="whole embryo"
/dev stage="shield stage, 6 hrs post-fertilisation"
/lab_host="E.coli XL1 blue MRF"
/clone_lib="zebrafish shield stage whole embryo cDNA
library MPMGP637"
/note="Vector: pSPort1; Site 1: NotI; Site 2: SalI;
oligo-dT-NotI primed, SalI adaptors, directionally cloned,
library normalised by oligonucleotide fingerprinting"

ORIGIN

Query Match 77.6%; Score 19.4; DB 3; Length 802;
Best Local Similarity 95.2%; Pred. No. 5.4e+02;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AAAAACTGCATCCCAATGTT 22

Db 293 AGAAATGTCATCCCAATGTT 273

RESULT 40

EX188572

LOCUS
EX188572
Danio rerio genomic clone DKEX-197G14, genomic survey sequence.

ACCESSION
EX188572

VERSION
EX188572.1 GI:28020458

KEYWORDS
GSS

SOURCE
Danio rerio (zebrafish)

ORGANISM

Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.

REFERENCE
1 (bases 1 to 807)

AUTHORS
Humphray, S.J., Huckle, E. and Durham, J.L.

TITLE
Direct Submission

JOURNAL
Submitted (27-JAN-2003) The Sanger Institute, Wellcome Trust Genome
Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries:

humquerry@sanger.ac.uk Unpublished

COMMENT
This sequence was generated from the T7 end of BAC 197G14. 197G14
is part of the Daniokey BAC Library created by R. Plasterk and N.V.
Keygene. Further details:
http://www.sanger.ac.uk/Projects/D_rerio/.

FEATURES

Location/Qualifiers
1..807
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="DKEX-197G14"
/tissue type="Testis"
/note="Vector pIndigoBAC-536"

ORIGIN

Query Match 77.6%; Score 19.4; DB 10; Length 807;
Best Local Similarity 95.2%; Pred. No. 5.4e+02;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AAAAACTGCATCCCAATGTT 22

Db 534 AGAAATGTCATCCCAATGTT 554

RESULT 41

CD169592

LOCUS

DEFINITION

MM1-0024T-M079-F09-U.G MM1-0024 Schistosoma mansoni cDNA clone

ACCESSION

VERSION
CD169592

KEYWORDS
EST.

SOURCE
CD169592.1 GI:34706196

ORGANISM

Schistosoma mansoni

Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea;

Strigeidida; Schistosomatidae; Schistosomatidae; Schistosoma.

REFERENCE
1 (bases 1 to 375)

AUTHORS

Verjovski-Almeida, S., DeMarco, R., Martins, E.A.L., Guimaraes, P.E.M.,

Ojopi, E.P.B., Paquola, A.C.M., Piazza, J.P., Nishiyama, M.Y. Jr.,

Kitajima, J.P., Adamson, R.E., Ashton, P.D., Bonaldo, M.F.,

Coulson, P.S., Dillon, G.P., Farias, L.P., Gregorio, S.P., Ho, P.L.,

Leite, R.A., Malaquias, L.C.C., Marques, R.C.P., Miyasato, P.A.,

Nascimento, A.L.T.O., Ohlweiler, F.P., Reis, E.M., Ribeiro, M.A.,

Sa, R.G., Stukart, G.C., Soares, M.B., Gargioni, C., Kawano, T.,

Rodrigues, V., Madeira, A.M.B.N., Wilson, R.A., Menck, C.F.M.,

Setubal, J.C., Leite, L.C.C. and Dias-Neto, E.

Transcriptome analysis of the acoelomate human parasite Schistosoma

mansoni

JOURNAL
Nat. Genet. 35 (2), 148-157 (2003)

PUBMED
12973350

COMMENT
Contact: Dr. Sergio Verjovski-Almeida

Departamento de Bioquímica

Instituto de Química - Universidade de São Paulo

Av. Prof. Lineu Prestes 748 sala 1200, 05508-900 São Paulo - SP,

Brasil

Tel: +55-11-3091-2173

Fax: +55-11-3091-2186

Email: verjo@iq.usp.br

This sequence was derived from the FAPESP Schistosoma mansoni EST

Genome Project. All sequences in the project were assembled and

annotated. This entry and all the assembled sequences can be seen

in the following URL <http://bioinfo.iq.usp.br/schisto/>

Plate: MM1-0024T-M079 row: 9 column: F.

FEATURES

Location/Qualifiers

1..375

/organism="Schistosoma mansoni"

/mol_type="mRNA"

/db_xref="taxon:6183"

/clone="MM1-0024T-M079-F09.G"

/sex="male"

/dev_stage="adult"

/lab_host="Mesocricetus auratus"

/clone_lib="MM1-0024"

ORIGIN

Query Match 76.8%; Score 19.2; DB 6; Length 375;

Best Local Similarity 87.5%; Pred. No. 5.8e+02;

Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAACTGCATCCCAATGTTAT 24

Db 348 AAAAACTGCATCCCAATGTTAT 371

RESULT 42

H70808/c

LOCUS

H70808

DEFINITION

IMAGE:214078 5', mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

REFERENCE

AUTHORS

1 (bases 1 to 403)

Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiapelli, B.,

CD169592
MM1-0024T-M079-F09-U.G MM1-0024 Schistosoma mansoni cDNA clone
MM1-0024T-M079-F09.G, mRNA sequence.

CD169592
CD169592.1 GI:34706196
EST.

Schistosoma mansoni
Schistosoma mansoni

Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea;
Strigeidida; Schistosomatidae; Schistosomatidae; Schistosoma.

REFERENCE
1 (bases 1 to 375)

AUTHORS

Verjovski-Almeida, S., DeMarco, R., Martins, E.A.L., Guimaraes, P.E.M.,

Ojopi, E.P.B., Paquola, A.C.M., Piazza, J.P., Nishiyama, M.Y. Jr.,

Kitajima, J.P., Adamson, R.E., Ashton, P.D., Bonaldo, M.F.,

Coulson, P.S., Dillon, G.P., Farias, L.P., Gregorio, S.P., Ho, P.L.,

Leite, R.A., Malaquias, L.C.C., Marques, R.C.P., Miyasato, P.A.,

Nascimento, A.L.T.O., Ohlweiler, F.P., Reis, E.M., Ribeiro, M.A.,

Sa, R.G., Stukart, G.C., Soares, M.B., Gargioni, C., Kawano, T.,

Rodrigues, V., Madeira, A.M.B.N., Wilson, R.A., Menck, C.F.M.,

Setubal, J.C., Leite, L.C.C. and Dias-Neto, E.

Transcriptome analysis of the acoelomate human parasite Schistosoma

mansoni

JOURNAL
Nat. Genet. 35 (2), 148-157 (2003)

PUBMED
12973350

COMMENT
Contact: Dr. Sergio Verjovski-Almeida

Departamento de Bioquímica

Instituto de Química - Universidade de São Paulo

Av. Prof. Lineu Prestes 748 sala 1200, 05508-900 São Paulo - SP,

Brasil

Tel: +55-11-3091-2173

Fax: +55-11-3091-2186

Email: verjo@iq.usp.br

This sequence was derived from the FAPESP Schistosoma mansoni EST

Genome Project. All sequences in the project were assembled and

annotated. This entry and all the assembled sequences can be seen

in the following URL <http://bioinfo.iq.usp.br/schisto/>

Plate: MM1-0024T-M079 row: 9 column: F.

FEATURES

Location/Qualifiers

1..375

/organism="Schistosoma mansoni"

/mol_type="mRNA"

/db_xref="taxon:6183"

/clone="MM1-0024T-M079-F09.G"

/sex="male"

/dev_stage="adult"

/lab_host="Mesocricetus auratus"

/clone_lib="MM1-0024"

ORIGIN

Query Match 76.8%; Score 19.2; DB 6; Length 375;

Best Local Similarity 87.5%; Pred. No. 5.8e+02;

Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAACTGCATCCCAATGTTAT 24

Db 348 AAAAACTGCATCCCAATGTTAT 371

RESULT 43

H70808

LOCUS

H70808

DEFINITION

IMAGE:214078 5', mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

REFERENCE

AUTHORS

1 (bases 1 to 403)

Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiapelli, B.,


```

JOURNAL Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. http://www.sanger.ac.uk/MICER
FEATURES
source      Location/Qualifiers
            1..649
            /organism="Mus musculus"
            /mol_type="genomic DNA"
            /db_xref="taxon:10090"
            /clone="MHPN379119"
            /clone_lib="MHPN"

ORIGIN
Query Match      76.8%; Score 19.2; DB 11; Length 649;
Best Local Similarity 87.5%; Pred. No. 6.4e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAAGTCATCCCAATGTTAT 24
    |||||
Db 73 AAAAAAAGTCATCCCAATGTAAT 96
    |||||

RESULT 45
CZ789932      711 bp DNA linear GSS 26-JUL-2005
LOCUS      OC_Ba0155L07.f OC_Ba Oryza coarctata genomic clone OC_Ba0155L07
DEFINITION      5', genomic survey sequence.
ACCESSION      CZ789932
VERSION      CZ789932.1 GI:71229785
KEYWORDS      GSS.
SOURCE      Oryza coarctata (Porteresia coarctata)
ORGANISM      Oryza coarctata
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE      1. (Bases 1 to 711)
AUTHORS      Kim,H., Collura,K., Wissotski,M., Byrne,M., Stum,D., Smart,D.,
            Rao,K., Luo,M., Jetty,R., Kudrna,D., Muller,C., Soderlund,C. and
            Wing,R.
TITLE      OMAP (Oryza Map Alignment Project)- Arizona Genomics Institute
JOURNAL      Unpublished (2005)
COMMENT      Contact: Rod A. Wing
            Arizona Genomics Institute
            University of Arizona
            Forbes Building Room 303, Tucson, AZ 85721-0036, USA
            Tel: 520 626 9595
            Fax: 520 621 1259
            Email: rwing@genome.arizona.edu
            PCR Primers
            FORWARD: TAA TAC GAC TCA CTA TAG GG
            BACKWARD: CAC TCA TTA GGC ACC CCA
            Plate: 0155 row: L column: 07
            Seq primer: TAA TAC GAC TCA CTA TAG GG
            Class: BAC ends.

FEATURES
source      Location/Qualifiers
            1..711
            /organism="Oryza coarctata"
            /mol_type="genomic DNA"
            /db_xref="taxon:77588"
            /clone="OC_Ba0155L07"
            /tissue type="leaves"
            /dev stage="mature"
            /lab_host="DH10B"
            /clone_lib="OC_Ba"
            /note="vector: pAGIBAC1; Site_1: HindIII; Site_2: HindIII"

ORIGIN
Query Match      76.8%; Score 19.2; DB 10; Length 711;
Best Local Similarity 87.5%; Pred. No. 6.5e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAAGTCATCCCAATGTTAT 24
    |||||
Db 687 AAAAAATGGCATCCCAATGTGAT 710
    |||||

```

AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
JOURNAL Tumor Gene Index
COMMENT Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Drs. Donald Brown and Liguang Cai
CDNA Library Preparation: CLONTECH
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLC3123 row: 1 column: 24
High quality sequence start: 169
High quality sequence stop: 516.
Features Location/Qualifiers
1. .1052
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone="IMAGE:6880393"
/dev_stage="metamorphosis stage 53"
/clone_lib="NICHG XGC Tadi"
/note="Organ: Developing Tadpole; Vector: pDNR-LIB;
Site 1: Sfi; Site 2: Sfi; 5' and 3' adaptors were used in
cloning as follows: 5' adaptor sequence:
5'-CAGCGCATATGGCC-3' and 3' adaptor sequence:
5'-ATTCTAGAGCGGAGCGGCACATG-dt(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.6 kb
(range 0.9-3.0 kb). 15/15 colonies contained inserts by
PCR. This library was enriched for full-length clones and
was constructed by Clontech Laboratories (Palo Alto, CA)."

ORIGIN
Query Match 75.2%; Score 18.8; DB 7; Length 418;
Best Local Similarity 90.9%; Pred. No. 9.1e+02;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 AAAAACTGCATCCCAATGTT 22
|||||
DB 233 AAAAACTGCTTCCCAATTT 212
|||||

RESULT 49
CV041200/c
LOCUS
DEFINITION 4138919 BARC_3CAL chicken mixed tissue Gallus gallus cDNA clone
3CAL 29009 5', mRNA sequence.
ACCESSION CV041200.1 GI:53560639
VERSION EST.
KEYWORDS Gallus gallus (chicken)
SOURCE Gallus gallus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 431)
AUTHORS Evock-Clover,C.M., Ashwell,C.M., McMurtry,J.P., Lillehoj,H.S.,
Matukumalli,L.K. and Van Tassel,C.P.
TITLE Characterization of expressed sequence tags generated from multiple
chicken tissues
JOURNAL Unpublished (2004)
COMMENT Contact: Christina M. Clover
Growth Biology Laboratory
Animal and Natural Resources Institute
Bldg. 200 Rm215 BARC-East, Beltsville, MD 20705, USA
Tel: 3015048224
Fax: 3015048623
Email: chrisclover@nri.barc.usda.gov
Single pass sequencing. Bases called and trimmed with phred
0.000925 using options -trim alt '' -trim fastavector identified by
cross match using options -minmatch 12 -minscore 12
Plate: 29 row: 0 column: 09
Seq primer: CCACATCAGCAGCGTGTGTAACCG
High quality sequence stop: 431.
Features Location/Qualifiers
1. .431
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="Leghorn and broiler"
/db_xref="taxon:9031"
/clone="3CAL 29009"
/lab_host="DH5alpha"

AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
JOURNAL Tumor Gene Index
COMMENT Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Drs. Donald Brown and Liguang Cai
CDNA Library Preparation: CLONTECH
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLC3123 row: 1 column: 24
High quality sequence start: 169
High quality sequence stop: 516.
Features Location/Qualifiers
1. .1052
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone="IMAGE:6880393"
/dev_stage="metamorphosis stage 53"
/clone_lib="NICHG XGC Tadi"
/note="Organ: Developing Tadpole; Vector: pDNR-LIB;
Site 1: Sfi; Site 2: Sfi; 5' and 3' adaptors were used in
cloning as follows: 5' adaptor sequence:
5'-CAGCGCATATGGCC-3' and 3' adaptor sequence:
5'-ATTCTAGAGCGGAGCGGCACATG-dt(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.6 kb
(range 0.9-3.0 kb). 15/15 colonies contained inserts by
PCR. This library was enriched for full-length clones and
was constructed by Clontech Laboratories (Palo Alto, CA)."

ORIGIN
Query Match 76.8%; Score 19.2; DB 6; Length 1052;
Best Local Similarity 87.5%; Pred. No. 7.1e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 AAAAACTGCATCCCAATGTTAT 24
|||||
DB 96 AAAAACTGCATACAGAATGTTAT 73
|||||

RESULT 48
CO503266/c
LOCUS
DEFINITION C0503266 418 bp mRNA linear EST 13-JUL-2004
clone GGEZCB1001D01.g chicken breast muscle - CB1 Gallus gallus cDNA
C0503266
C0503266.1 GI:50273452
EST.
Gallus gallus (chicken)
Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 418)
AUTHORS Alves,H.J., Jorge,E.C., Marchesin,M.L., Monteiro-Vitorello,C.B.,
Patricio,M., Ledur,M.C. and Coutinho,L.L.
TITLE Discovery of new genes expressed in the chicken breast muscle
JOURNAL Unpublished (2004)
COMMENT Contact: Helena J. Alves
Laboratory of Animal Biotechnology, Dep. of Animal Production
ESALQ - University of Sao Paulo
Av. Pádua Dias, 11, Piracicaba, SP, 13418-900, Brazil
Tel: 55 19 3429 4434
Fax: 55 19 3429 4285
Email: hjalves@esalq.usp.br and llcoutin@esalq.usp.br
PCR Primers
BACKWARD: T7.
Features Location/Qualifiers
1. .418
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="Leghorn and broiler"
/db_xref="taxon:9031"
/clone="3CAL 29009"
/lab_host="DH5alpha"

/clone lib="BARC 3GAL chicken mixed tissue"
/note="Vector: pBluescript, SK+, Stratagene; Site 1: NotI;
Site 2: EcoRI; Normalized library of pooled RNA isolated
from whole brain, ultimobranchial gland, parathyroid
gland, cecal tonsil and primordial germ cells Multiple"

ORIGIN

Query Match 75.2%; Score 18.8; DB 7; Length 431;
Best Local Similarity 90.9%; Pred. No. 9.1e+02; Mismatches 2; Indels 0; Gaps 0;
Matches 20; Conservative 0;

Qy 1 AAAAACTGCATCCCAATGTT 22
|||||
Db 206 AAAAACTGCTTCCCAATTT 185

RESULT 50
BU360170/c
LOCUS BU360170 443 bp mRNA linear EST 28-NOV-2002
DEFINITION 603477364F1 CSEQCHN71 Gallus gallus cDNA clone ChEST363015 5', mRNA
sequence.
ACCESSION BU360170
VERSION BU360170.1 GI:25968171
KEYWORDS EST.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 443)
Boardman, P. E., Sanz-Ezquerro, J., Overton, I. M., Burt, D. W., Bosch, E.,
Fong, W. T., Tickle, C., Brown, W. R. A., Wilson, S. A. and Hubbard, S. J.
A Comprehensive Collection of Chicken cDNAs
Curr. Biol. 12 (22), 1965-1969 (2002)
12445392

COMMENT
Contact: Simon Hubbard
Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology
(UMIST)

PO Box 88, Manchester, M60 1QD, UK
Tel: 01612008930
Fax: 01612360409
Email: Simon.Hubbard@umist.ac.uk.

FEATURES

source

1..443
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="White Leghorn, Hisex"
/db_xref="taxon:9031"
/clone="ChEST363015"
/dev_stage="36"
/lab_host="DH10B"
/clone_lib="CSEQCHN71"

/note="Organ: hearts; Vector: pBluescript II KS(+);
Site 1: EcoRI; Site 2: NotI. This normalized library was
constructed from 1 million independent clones. cDNA
synthesis was initiated using an oligo(dT) primer, using
methylated C in the first strand synthesis reaction.
Following this first strand reaction, double-stranded cDNA
was blunted, ligated to NotI adapters, digested with
EcoRI, size-selected, and cloned into the NotI and EcoRI
compatible sites of a custom modified MCS of the
pBluescript (KS+) vector. The library was normalized in 2
rounds using conditions adapted from Soares et al., PNAS
(1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
(1996): 791, except that a significantly longer
reannealing hybridization was used."

ORIGIN

Query Match 75.2%; Score 18.8; DB 5; Length 443;
Best Local Similarity 90.9%; Pred. No. 9.2e+02;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AAAAACTGCATCCCAATGTT 22
|||||
Db 177 AAAAACTGCTTCCCAATTT 156

Search completed: February 3, 2006, 22:02:31
Job time : 2961.67 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 3, 2006, 11:35:31 ; Search time 80.6667 Seconds
(without alignments)
550.897 Million cell updates/sec

Title: US-10-719-900-25

Perfect score: 25

Sequence: 1 aaaaacgcatcccaatgttatg 25

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database :

Issued Patents NA:
1: /cgn2_6/ptodata/1/ina/1.COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5.COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq.*
5: /cgn2_6/ptodata/1/ina/H.COMB.seq.*
6: /cgn2_6/ptodata/1/ina/PCTUS.COMB.seq.*
7: /cgn2_6/ptodata/1/ina/PP.COMB.seq.*
8: /cgn2_6/ptodata/1/ina/RE.COMB.seq.*
9: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	21.8	87.2	270	3	US-09-016-434-1010
2	21.8	87.2	460	3	US-09-621-376-1487
3	20.4	81.6	1917	3	US-09-710-279-1467
4	20.4	81.6	1953	3	US-09-134-001C-1607
5	20.4	81.6	3276	3	US-09-710-279-4240
6	19.2	76.8	601	3	US-09-949-016-179347
7	19.2	76.8	263693	3	US-09-949-016-12386
8	19.2	76.8	263694	3	US-09-949-016-16915
9	18.6	74.4	168174	3	US-10-071-411A-63
10	18.6	74.4	168273	3	US-10-071-411A-2
11	17.8	71.2	300598	3	US-09-949-016-11868
12	17.8	71.2	302604	3	US-09-949-016-14588
13	17.8	71.2	302604	3	US-09-949-016-14589
14	17.8	71.2	308362	3	US-09-949-016-17119
15	17.6	70.4	601	3	US-09-949-016-62866
16	17.6	70.4	1770	3	US-09-489-039A-109
17	17.6	70.4	1803	3	US-09-489-039A-209
18	17.6	70.4	12563	3	US-09-949-016-4982
19	17.6	70.4	103792	3	US-09-949-016-13553
20	17.6	70.4	103993	3	US-09-949-016-12317
21	17.6	70.4	138693	3	US-09-949-016-16724
22	17.6	70.4	285986	3	US-09-949-016-12287
23	17.6	70.4	288031	3	US-09-949-016-14864
24	17.4	69.6	2179	3	US-09-620-312D-78
					Sequence 1010, Ap
					Sequence 1487, A
					Sequence 1467, Ap
					Sequence 1607, Ap
					Sequence 4240, Ap
					Sequence 179347, A
					Sequence 12386, A
					Sequence 16915, A
					Sequence 63, Appl
					Sequence 2, Appl
					Sequence 11868, A
					Sequence 14588, A
					Sequence 14589, A
					Sequence 17119, A
					Sequence 62866, A
					Sequence 109, App
					Sequence 209, App
					Sequence 4982, Ap
					Sequence 13553, A
					Sequence 12317, A
					Sequence 16724, A
					Sequence 12287, A
					Sequence 14864, A
					Sequence 78, Appl

25	17.2	68.8	881	3	US-09-270-767-11666	Sequence 11666, A
26	17.2	68.8	3198	3	US-09-345-236B-83	Sequence 83, Appl
27	17.2	68.8	3198	3	US-09-345-236B-87	Sequence 87, Appl
28	17.2	68.8	3198	3	US-09-345-236B-90	Sequence 90, Appl
29	17.2	68.8	3198	3	US-09-345-236B-92	Sequence 92, Appl
30	17.2	68.8	3918	3	US-08-836-323-1	Sequence 1, Appl
31	17.2	68.8	22605	3	US-09-949-016-17001	Sequence 17001, A
32	17.2	68.8	40655	3	US-09-949-016-12032	Sequence 12032, A
33	17.2	68.8	40655	3	US-09-949-016-15919	Sequence 15919, A
34	17.2	68.8	41696	3	US-09-949-016-12444	Sequence 12444, A
35	17.2	68.8	58782	3	US-09-949-016-16851	Sequence 16851, A
36	17.2	68.8	94142	3	US-09-949-016-18553	Sequence 16553, A
37	17.2	68.8	157822	3	US-09-949-016-16723	Sequence 16723, A
38	17	68.0	134	3	US-09-513-999C-16990	Sequence 16990, A
39	17	68.0	171	3	US-09-513-999C-22791	Sequence 22791, A
40	17	68.0	173	3	US-09-270-767-6867	Sequence 6867, Ap
41	17	68.0	173	3	US-09-270-767-22149	Sequence 22149, A
42	17	68.0	243	3	US-09-134-000C-150	Sequence 150, App
43	17	68.0	430	3	US-09-270-767-4566	Sequence 4566, Ap
44	17	68.0	430	3	US-09-270-767-19848	Sequence 19848, A
45	17	68.0	601	3	US-09-949-016-131417	Sequence 131417, Ap
46	17	68.0	654	3	US-09-543-681A-3477	Sequence 3477, Ap
47	17	68.0	918	3	US-09-328-352-2668	Sequence 2668, Ap
48	17	68.0	20229	3	US-09-949-016-16649	Sequence 16649, A
49	17	68.0	76810	3	US-09-949-016-12528	Sequence 12528, A
50	17	68.0	107820	3	US-09-792-616-1	Sequence 1, Appl
51	17	68.0	124700	3	US-09-949-016-11817	Sequence 11817, A
52	17	68.0	124701	3	US-09-949-016-15439	Sequence 15439, A
53	17	68.0	135171	3	US-09-949-016-15617	Sequence 15617, A
54	17	68.0	150409	3	US-09-949-016-12290	Sequence 12290, A
55	17	68.0	150409	3	US-09-949-016-12938	Sequence 12938, A
56	17	68.0	786431	3	US-09-751-389-3	Sequence 3, Appl
57	16.8	67.2	245	3	US-09-513-999C-8833	Sequence 8833, Ap
58	16.8	67.2	601	3	US-09-949-016-83336	Sequence 8336, A
59	16.8	67.2	1200	3	US-09-602-472A-11	Sequence 11, Appl
60	16.8	67.2	3456	3	US-09-602-472A-12	Sequence 12, Appl
61	16.8	67.2	143550	3	US-09-949-016-14143	Sequence 14143, A
62	16.8	67.2	168575	3	US-09-426-290-1	Sequence 1, Appl
63	16.8	67.2	392000	3	US-10-027-983-11	Sequence 11, Appl
64	16.6	66.4	380	3	US-09-347-798-5	Sequence 5, Appl
65	16.6	66.4	426	3	US-09-270-767-5056	Sequence 5056, Ap
66	16.6	66.4	426	3	US-09-270-767-20338	Sequence 20338, A
67	16.6	66.4	474	3	US-09-489-039A-7044	Sequence 7044, Ap
68	16.6	66.4	601	3	US-09-949-016-55241	Sequence 55241, A
69	16.6	66.4	601	3	US-09-949-016-74666	Sequence 74666, A
70	16.6	66.4	601	3	US-09-949-016-74865	Sequence 74865, A
71	16.6	66.4	601	3	US-09-949-016-75064	Sequence 75064, A
72	16.6	66.4	601	3	US-09-949-016-75263	Sequence 75263, A
73	16.6	66.4	601	3	US-09-949-016-75462	Sequence 75462, A
74	16.6	66.4	601	3	US-09-949-016-77347	Sequence 77347, A
75	16.6	66.4	601	3	US-09-949-016-77348	Sequence 77348, A
76	16.6	66.4	601	3	US-09-949-016-106020	Sequence 106020, A
77	16.6	66.4	601	3	US-09-949-016-106219	Sequence 106219, A
78	16.6	66.4	601	3	US-09-949-016-106418	Sequence 106418, A
79	16.6	66.4	601	3	US-09-949-016-106617	Sequence 106617, A
80	16.6	66.4	601	3	US-09-949-016-106816	Sequence 106816, A
81	16.6	66.4	601	3	US-09-949-016-128729	Sequence 128729, A
82	16.6	66.4	601	3	US-09-949-016-153399	Sequence 153399, A
83	16.6	66.4	601	3	US-09-949-016-153400	Sequence 153400, A
84	16.6	66.4	601	3	US-09-949-016-153401	Sequence 153401, A
85	16.6	66.4	601	3	US-09-949-016-153402	Sequence 153402, A
86	16.6	66.4	770	3	US-09-504-358-13	Sequence 1259, Ap
87	16.6	66.4	771	3	US-09-504-358-13	Sequence 13, Appl
88	16.6	66.4	771	3	US-09-954-314-13	Sequence 13, Appl
89	16.6	66.4	771	3	US-10-230-562-13	Sequence 9, Appl
90	16.6	66.4	1342	3	US-09-500-569-9	Sequence 9, Appl
91	16.6	66.4	1342	3	US-09-971-823B-94	Sequence 94, Appl
92	16.6	66.4	2161	3	US-09-509-712B-99	Sequence 1111, Ap
93	16.6	66.4	2439	2	US-09-489-039A-7111	Sequence 1, Appl
94	16.6	66.4	2814	6	US-07-781-034-1	Sequence 1, Appl
95	16.6	66.4	2814	6	PCT-US92-08328-1	Sequence 1, Appl
96	16.6	66.4	9848	3	US-09-385-222A-3	Sequence 3, Appl
97	16.6	66.4	10629	3	US-09-504-358-15	Sequence 15, Appl

Sequence 15, Appl
Sequence 15, Appl
Sequence 12071, A
Sequence 12673, A
Sequence 16172, A
Sequence 16173, A
Sequence 16174, A
Sequence 16175, A
Sequence 12902, A
Sequence 14386, A
Sequence 17433, A
Sequence 13080, A
Sequence 16329, A
Sequence 13353, A
Sequence 12273, A
Sequence 16050, A
Sequence 14000, A
Sequence 15393, A
Sequence 13781, A
Sequence 13922, A
Sequence 13923, A
Sequence 13924, A
Sequence 13925, A
Sequence 13926, A
Sequence 14699, A
Sequence 14700, A
Sequence 14701, A
Sequence 14702, A
Sequence 14703, A
Sequence 1, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 981, App
Sequence 14421, A
Sequence 11929, A
Sequence 15421, A
Sequence 12822, A
Sequence 15524, A
Sequence 17532, A
Sequence 59, Appl
Sequence 2450, Ap
Sequence 71499, A
Sequence 71500, A
Sequence 91914, A
Sequence 165425, A
Sequence 165482, A
Sequence 194855, A
Sequence 685, App
Sequence 686, App
Sequence 6221, Ap
Sequence 10285, A

ALIGNMENTS

RESULT 1
US-09-016-434-1010
; Sequence 1010, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA

COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/016.434
FILING DATE: HEREWITH
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0002 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 1010:
SEQUENCE CHARACTERISTICS:
LENGTH: 270 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: BRAINOT04
CLONE: 926642
US-09-016-434-1010

Query Match 87.2%; Score 21.8; DB 3; Length 270;
Best Local Similarity 92.0%; Pred. No. 0.46;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAACTGCATCCCAATGTTATG 25
Db 164 AAAAACTGCATCCCAATGTTATG 188

RESULT 2
US-09-621-976-14487
; Sequence 14487, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621.976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 14487
; LENGTH: 460
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-14487

Query Match 87.2%; Score 21.8; DB 3; Length 460;
Best Local Similarity 92.0%; Pred. No. 0.51;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAACTGCATCCCAATGTTATG 25
Db 136 AAAAACTGCATCCCAATGTTATG 160

RESULT 3
US-09-710-279-1467

```
; Sequence 1467, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/09/710,279
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1467
; LENGTH: 1917
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-09-710-279-1467

Query Match      81.6%; Score 20.4; DB 3; Length 1917;
Best Local Similarity 95.5%; Pred. No. 3.5;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      3 AAAAAGTCATCCCAATGTTAT 24
      |||||
Db      1441 AAAAAGTCATCCCAATGTTAT 1462

RESULT 4
US-09-134-001C-1607
; Sequence 1607, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 1607
; LENGTH: 1953
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-1607

Query Match      81.6%; Score 20.4; DB 3; Length 1953;
Best Local Similarity 95.5%; Pred. No. 3.5;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      3 AAAAAGTCATCCCAATGTTAT 24
      |||||
Db      1477 AAAAAGTCATCCCAATGTTAT 1498

RESULT 5
US-09-710-279-4240
; Sequence 4240, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/09/710,279
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
```

```
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4240
; LENGTH: 3276
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-09-710-279-4240

Query Match      81.6%; Score 20.4; DB 3; Length 3276;
Best Local Similarity 95.5%; Pred. No. 3.9;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      3 AAAAAGTCATCCCAATGTTAT 24
      |||||
Db      1193 AAAAAGTCATCCCAATGTTAT 1214

RESULT 6
US-09-949-016-179347
; Sequence 179347, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 179347
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-179347

Query Match      76.8%; Score 19.2; DB 3; Length 601;
Best Local Similarity 87.5%; Pred. No. 10;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      2 AAAAAGTCATCCCAATGTTATG 25
      |||||
Db      190 AAAAAGTCATCCCAATGTTATG 213

RESULT 7
US-09-949-016-12386
; Sequence 12386, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 12386
; LENGTH: 263693
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12386

Query Match          76.8%; Score 19.2; DB 3; Length 263693;
Best Local Similarity 87.5%; Pred. No. 42;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAACTGCATCCCAATGTTATG 25
   ||||| ||||| ||||| ||||| |||||
Db 115551 AAAAACTGCATATCCAGTGTATG 115574

RESULT 8
US-09-949-016-16915
; Sequence 63, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16915
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-16915

Query Match          76.8%; Score 19.2; DB 3; Length 263694;
Best Local Similarity 87.5%; Pred. No. 42;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAACTGCATCCCAATGTTATG 25
   ||||| ||||| ||||| ||||| |||||
Db 115551 AAAAACTGCATATCCAGTGTATG 115574

RESULT 9
US-10-071-411A-63
; Sequence 63, Application US/10071411A
; Patent No. 6797475
; GENERAL INFORMATION:
; APPLICANT: Glenn Barnes
; APPLICANT: Joanne Meyer
; TITLE OF INVENTION: Detection of Polymorphisms in the Human
; FILE REFERENCE: 5-Lipoxygenase Gene
; CURRENT APPLICATION NUMBER: US/10/071,411A
; PRIOR FILING DATE: 2002-02-07
; PRIOR APPLICATION NUMBER: 60/267,515
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: 60/314,248
; PRIOR FILING DATE: 2001-08-21
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 63
; LENGTH: 168174
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
```

```
; LOCATION: (1)...(168174)
; OTHER INFORMATION: n = A,T,C or G
US-10-071-411A-63

Query Match          74.4%; Score 18.6; DB 3; Length 168174;
Best Local Similarity 84.0%; Pred. No. 74;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAACTGCATCCCAATGTTATG 25
   ||||| ||||| ||||| ||||| |||||
Db 91147 AAAGAACTGCATCTTAATGTTCTG 91171

RESULT 10
US-10-071-411A-2
; Sequence 2, Application US/10071411A
; Patent No. 6797475
; GENERAL INFORMATION:
; APPLICANT: Glenn Barnes
; APPLICANT: Joanne Meyer
; TITLE OF INVENTION: Detection of Polymorphisms in the Human
; FILE REFERENCE: MRI-021
; CURRENT APPLICATION NUMBER: US/10/071,411A
; PRIOR FILING DATE: 2002-02-07
; PRIOR APPLICATION NUMBER: 60/267,515
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: 60/314,248
; PRIOR FILING DATE: 2001-08-21
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 168273
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1) -(168273)
; OTHER INFORMATION: n = A,T,C or G
US-10-071-411A-2

Query Match          74.4%; Score 18.6; DB 3; Length 168273;
Best Local Similarity 84.0%; Pred. No. 74;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAACTGCATCCCAATGTTATG 25
   ||||| ||||| ||||| ||||| |||||
Db 91246 AAAGAACTGCATCTTAATGTTCTG 91270

RESULT 11
US-09-949-016-11868
; Sequence 11868, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11868
; LENGTH: 300598
; TYPE: DNA
; ORGANISM: Human
```

```
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(300598)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-11868

Query Match          71.2%; Score 17.8; DB 3; Length 300598;
Best Local Similarity 90.5%; Pred. No. 2.1e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 AACTGCATCCCCCAATGTTATG 25
   ||||||| |||||||
Db 205760 AATTGCATCCACAATGTTATG 205780

RESULT 12
US-09-949-016-14588
; Sequence 14588, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14588
; LENGTH: 302604
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(302604)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14588

Query Match          71.2%; Score 17.8; DB 3; Length 302604;
Best Local Similarity 90.5%; Pred. No. 2.1e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 AACTGCATCCCCCAATGTTATG 25
   ||||||| |||||||
Db 255766 AATTGCATCCACAATGTTATG 255786

RESULT 13
US-09-949-016-14589
; Sequence 14589, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14589

Query Match          71.2%; Score 17.8; DB 3; Length 302604;
Best Local Similarity 90.5%; Pred. No. 2.1e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 AACTGCATCCCCCAATGTTATG 25
   ||||||| |||||||
Db 255766 AATTGCATCCACAATGTTATG 255786

RESULT 14
US-09-949-016-17119
; Sequence 17119, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17119
; LENGTH: 308362
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(308362)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17119

Query Match          71.2%; Score 17.8; DB 3; Length 308362;
Best Local Similarity 90.5%; Pred. No. 2.1e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 AACTGCATCCCCCAATGTTATG 25
   ||||||| |||||||
Db 255582 AATTGCATCCACAATGTTATG 255602

RESULT 15
US-09-949-016-62866
; Sequence 62866, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14589
```

```
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 62866
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-62866

Query Match          70.4%; Score 17.6; DB 3; Length 601;
Best Local Similarity 83.3%; Pred. No. 64;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAACTGCATCCCAATGTTAT 24
    ||||| ||||| ||||| ||||| |||||
Db 199 AAAAAATGCATACCAATGGTAT 222

RESULT 16
US-09-489-039A-109/c
; Sequence 109, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 109
; LENGTH: 1770
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-109

Query Match          70.4%; Score 17.6; DB 3; Length 1770;
Best Local Similarity 83.3%; Pred. No. 82;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AAAAACTGCATCCCAATGTTATG 25
    ||||| ||||| ||||| ||||| |||||
Db 1338 AAAAACTTCATCCCGAGTATG 1315

RESULT 17
US-09-489-039A-209
; Sequence 209, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 209
; LENGTH: 1803
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-209

Query Match          70.4%; Score 17.6; DB 3; Length 1803;
Best Local Similarity 83.3%; Pred. No. 82;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AAAAACTGCATCCCAATGTTATG 25
    ||||| ||||| ||||| ||||| |||||
Db 406 AAAAACTTCATCCCGAGTATG 429

RESULT 18
US-09-949-016-4982
; Sequence 4982, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4982
; LENGTH: 12563
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-4982

Query Match          70.4%; Score 17.6; DB 3; Length 12563;
Best Local Similarity 83.3%; Pred. No. 1.3e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAACTGCATCCCAATGTTAT 24
    ||||| ||||| ||||| ||||| |||||
Db 3191 AGAAAAAGACCATCCCAATGTTAT 3214

RESULT 19
US-09-949-016-13553/c
; Sequence 13553, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13553
; LENGTH: 103792
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(103792)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13553

Query Match          70.4%; Score 17.6; DB 3; Length 103792;
Best Local Similarity 83.3%; Pred. No. 2.1e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAACTGCATCCCAATGTTAT 24
    ||||| ||||| ||||| ||||| |||||
Db 54553 AAAAAAATGCATACCCCAATGGTAT 54530
```

RESULT 20
US-09-949-016-12317/c
; Sequence 12317, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12317
; LENGTH: 103993
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(103993)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12317

Query Match 70.4%; Score 17.6; DB 3; Length 103993;
Best Local Similarity 83.3%; Pred. No. 2.1e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAACTGCATCCCAATGTTAT 24
||||| ||||| ||||| ||||| |||||
DB 54553 AAAAAATGCATACCCATGTTAT 54530

RESULT 21
US-09-949-016-16724
; Sequence 16724, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16724
; LENGTH: 138693
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-16724

Query Match 70.4%; Score 17.6; DB 3; Length 138693;
Best Local Similarity 83.3%; Pred. No. 2.2e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAACTGCATCCCAATGTTAT 24
||||| ||||| ||||| ||||| |||||
DB 50685 AGAAGACCATCCCAATGTTAT 50708

RESULT 22
US-09-949-016-12287/c
; Sequence 12287, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12287
; LENGTH: 285986
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(285986)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12287

Query Match 70.4%; Score 17.6; DB 3; Length 285986;
Best Local Similarity 83.3%; Pred. No. 2.5e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAACTGCATCCCAATGTTAT 24
||||| ||||| ||||| ||||| |||||
DB 103371 AAAAACTCAATCTCCAATGTTAT 103348

RESULT 23
US-09-949-016-14864/c
; Sequence 14864, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14864
; LENGTH: 288031
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(288031)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14864

Query Match 70.4%; Score 17.6; DB 3; Length 288031;
Best Local Similarity 83.3%; Pred. No. 2.6e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAACTGCATCCCAATGTTAT 24
||||| ||||| ||||| ||||| |||||

```
Db 153416 AAGAACTCAATCTCAATGTTAT 153393

RESULT 24
US-09-620-312D-78/c
; Sequence 78, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-Hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aigong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunging
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; FILE REFERENCE: Polypeptides
; FILE REFERENCE: 784CIP28
; CURRENT APPLICATION NUMBER: US/09/620,312D
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 78
; LENGTH: 2179
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (284)..(1957)
US-09-620-312D-78

Query Match 69.6%; Score 17.4; DB 3; Length 2179;
Best Local Similarity 94.7%; Pred. No. 1.1e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 ACTGCATCCCAATGTTAT 24
Db 1272 ACTGCAACCCCAATGTTAT 1254

RESULT 25
US-09-270-767-11666
; Sequence 11666, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11666
; LENGTH: 881
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-11666

Query Match 68.8%; Score 17.2; DB 3; Length 881;

Db 153416 AAGAACTCAATCTCAATGTTAT 153393

Best Local Similarity 86.4%; Pred. No. 1.1e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 AAAACTGCATCCCAATGTTAT 24
Db 738 AAAACTGAATCAACAATGTTAT 759

RESULT 26
US-09-345-236B-83
; Sequence 83, Application US/09345236B
; Patent No. 6521454
; GENERAL INFORMATION:
; APPLICANT: Becnel, James J.
; APPLICANT: Tuku, Fukuda
; APPLICANT: Moser, Bettina
; APPLICANT: Cockburn, Andrew
; APPLICANT: White, Susan E.
; APPLICANT: Undeen, Albert H.
; TITLE OF INVENTION: No. 6521454el Baculoviruses, Insecticidal
; FILE REFERENCE: 21042.0004
; CURRENT APPLICATION NUMBER: US/09/345,236B
; CURRENT FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 148
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 83
; LENGTH: 3198
; TYPE: DNA
; ORGANISM: mosquito baculovirus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2545)...(3198)
; NAME/KEY: CDS
; LOCATION: (742)...(1056)
; NAME/KEY: CDS
; LOCATION: (1837)...(1998)
US-09-345-236B-83

Query Match 68.8%; Score 17.2; DB 3; Length 3198;
Best Local Similarity 86.4%; Pred. No. 1.1e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 AAAACTGCATCCCAATGTTAT 24
Db 1279 AAAATTGCATCCCAATGTTAT 1300

RESULT 27
US-09-345-236B-87/c
; Sequence 87, Application US/09345236B
; Patent No. 6521454
; GENERAL INFORMATION:
; APPLICANT: Becnel, James J.
; APPLICANT: Tuku, Fukuda
; APPLICANT: Moser, Bettina
; APPLICANT: Cockburn, Andrew
; APPLICANT: White, Susan E.
; APPLICANT: Undeen, Albert H.
; TITLE OF INVENTION: No. 6521454el Baculoviruses, Insecticidal
; FILE REFERENCE: 21042.0004
; CURRENT APPLICATION NUMBER: US/09/345,236B
; CURRENT FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 148
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 87
; LENGTH: 3198
; TYPE: DNA
; ORGANISM: mosquito baculovirus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (428)...(3195)
```



```
; NAME/KEY: CDS
; LOCATION: (20)....(241)
US-09-345-236B-87

Query Match      68.8%; Score 17.2; DB 3; Length 3198;
Best Local Similarity 86.4%; Pred. No. 1.5e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 AAAAAGTCATCCCAATGTTAT 24
   ||||| ||||| ||||| |||||
Db 1920 AAAATTGCATCCCAATGTTAT 1899

RESULT 28
US-09-345-236B-90/c
; Sequence 90, Application US/09345236B
; Patent No. 6521454
; GENERAL INFORMATION:
; APPLICANT: Becnel, James J.
; APPLICANT: Tokuo, Fukuda
; APPLICANT: Moser, Bettina
; APPLICANT: Cockburn, Andrew
; APPLICANT: White, Susan E.
; APPLICANT: Undeen, Albert H.
; TITLE OF INVENTION: No. 6521454el Baculoviruses, Insecticidal
; FILE REFERENCE: 21042.0004
; CURRENT APPLICATION NUMBER: US/09/345,236B
; CURRENT FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 148
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 90
; LENGTH: 3198
; TYPE: DNA
; ORGANISM: mosquito baculovirus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (667)....(897)
US-09-345-236B-90

Query Match      68.8%; Score 17.2; DB 3; Length 3198;
Best Local Similarity 86.4%; Pred. No. 1.5e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 AAAAAGTCATCCCAATGTTAT 24
   ||||| ||||| ||||| |||||
Db 1920 AAAATTGCATCCCAATGTTAT 1899

RESULT 29
US-09-345-236B-92/c
; Sequence 92, Application US/09345236B
; Patent No. 6521454
; GENERAL INFORMATION:
; APPLICANT: Becnel, James J.
; APPLICANT: Tokuo, Fukuda
; APPLICANT: Moser, Bettina
; APPLICANT: Cockburn, Andrew
; APPLICANT: White, Susan E.
; APPLICANT: Undeen, Albert H.
; TITLE OF INVENTION: No. 6521454el Baculoviruses, Insecticidal
; FILE REFERENCE: 21042.0004
; CURRENT APPLICATION NUMBER: US/09/345,236B
; CURRENT FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 148
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 92
; LENGTH: 3198
; TYPE: DNA
; ORGANISM: mosquito baculovirus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (667)....(897)
US-09-345-236B-92

; LOCATION: (696)....(848)
US-09-345-236B-92

Query Match      68.8%; Score 17.2; DB 3; Length 3198;
Best Local Similarity 86.4%; Pred. No. 1.5e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 AAAAAGTCATCCCAATGTTAT 24
   ||||| ||||| ||||| |||||
Db 1920 AAAATTGCATCCCAATGTTAT 1899

RESULT 30
US-08-836-329-1
; Sequence 1, Application US/08836329
; Patent No. 6090546
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Method for the Detection of Ras Oncogenes,
; IN PARTICULAR The K-Ras Oncogene
; NUMBER OF SEQUENCES: 16
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/836,329
; FILING DATE:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3918 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: exon
; LOCATION: 825..947
; FEATURE:
; NAME/KEY: exon
; LOCATION: 1388..1567
; FEATURE:
; NAME/KEY: exon
; LOCATION: 2118..2278
; FEATURE:
; NAME/KEY: exon
; LOCATION: 3034..3158
; NAME/KEY: exon
; LOCATION: 3459..3616
; US-08-836-329-1

Query Match      68.8%; Score 17.2; DB 3; Length 3918;
Best Local Similarity 86.4%; Pred. No. 1.5e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 AAAAAGTCATCCCAATGTTAT 24
   ||||| ||||| ||||| |||||
Db 2879 AAAAAGTCATCCCAATGTTAT 2900

RESULT 31
US-09-949-016-17001
; Sequence 17001, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
```

; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17001
; LENGTH: 22605
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(22605)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17001

Query Match 68.8%; Score 17.2; DB 3; Length 22605;
Best Local Similarity 86.4%; Pred. No. 2.3e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 AAAAGTGCATCCCAATGTTAT 24
|||||
Db 13612 AAAAGTGCATCCTCAGTTTAT 13633

RESULT 32
US-09-949-016-12032/c
; Sequence 12032, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12032
; LENGTH: 40655
; TYPE: DNA
; ORGANISM: Human
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12032

Query Match 68.8%; Score 17.2; DB 3; Length 40655;
Best Local Similarity 86.4%; Pred. No. 2.6e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAGTGCATCCCAATGTT 22
|||||
Db 29026 AAAAGTGCATACCGAATGTT 29005

RESULT 33
US-09-949-016-15919/c
; Sequence 15919, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15919
; LENGTH: 40655
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15919

Query Match 68.8%; Score 17.2; DB 3; Length 40655;
Best Local Similarity 86.4%; Pred. No. 2.6e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAGTGCATCCCAATGTT 22
|||||
Db 29026 AAAAGTGCATACCGAATGTT 29005

RESULT 34
US-09-949-016-12444
; Sequence 12444, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12444
; LENGTH: 41696
; TYPE: DNA
; ORGANISM: Human
; NAME/KEY: misc_feature
; LOCATION: (1)...(41696)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12444

Query Match 68.8%; Score 17.2; DB 3; Length 41696;
Best Local Similarity 86.4%; Pred. No. 2.6e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 AAAAAGTGCATCCCAATGTTAT 24
|||||
Db 13612 AAAAGTGCATCCTCAGTTTAT 13633

RESULT 35
US-09-949-016-16851
; Sequence 16851, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16851
; LENGTH: 58782
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-16851

Query Match 68.8%; Score 17.2; DB 3; Length 58782;
Best Local Similarity 86.4%; Pred. No. 2.9e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAACTGCATCCCAATGTTA 23
||||| |||||||||
DB 47134 AAAAACTTAATCCCAATGTGA 47155

RESULT 36
US-09-949-016-16553/c
; Sequence 16553, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16553
; LENGTH: 94142
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(94142)
; OTHER INFORMATION: n = A, T, C or G
US-09-949-016-16553

Query Match 68.8%; Score 17.2; DB 3; Length 94142;
Best Local Similarity 86.4%; Pred. No. 3.2e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 AAATGTCATCCCAATGTTATG 25
||||| |||||||||
DB 55395 AAATTAATCCCAATGTGATG 55374

RESULT 37
US-09-949-016-16723
; Sequence 16723, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16723
; LENGTH: 157822
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-16723

Query Match 68.8%; Score 17.2; DB 3; Length 157822;
Best Local Similarity 86.4%; Pred. No. 3.5e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 AAAAATGTCATCCCAATGTTAT 24
||||| |||||||||
DB 86913 AATACTGCATCCCAATGTGAT 86934

RESULT 38
US-09-513-999C-16990/c
; Sequence 16990, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 16990
; LENGTH: 134
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 18
; OTHER INFORMATION: w=a or t
US-09-513-999C-16990

Query Match 68.0%; Score 17; DB 3; Length 134;
Best Local Similarity 80.0%; Pred. No. 90;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAAATGTCATCCCAATGTTATG 25
||||| |||||||||
DB 68 ATAAACTATACATCCCAATTTTIG 44

RESULT 39
US-09-513-999C-22791/c
; Sequence 22791, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26

```
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 22791
; LENGTH: 171
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 6
; OTHER INFORMATION: h=a or c or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 7
; OTHER INFORMATION: n=a, g, c or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 8
; OTHER INFORMATION: n=a, g, c or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 9
; OTHER INFORMATION: n=a, g, c or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 160
; OTHER INFORMATION: w=a or t
; OTHER INFORMATION: w=a or t
US-09-513-999C-22791

Query Match      68.0%; Score 17; DB 3; Length 171;
Best Local Similarity 80.0%; Pred. No. 95;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAACTGCATCCCAATGTTATG 25
   ||||| ||||| ||||| ||||| |||||
Db 41 AAAAAAATCTATCGCAATGTTATG 17

RESULT 40
US-09-270-767-6867/c
; Sequence 6867, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6867
; LENGTH: 173
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-6867

Query Match      68.0%; Score 17; DB 3; Length 173;
Best Local Similarity 80.0%; Pred. No. 95;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAACTGCATCCCAATGTTATG 25
   ||||| ||||| ||||| ||||| |||||
Db 112 AAACAACTGCAACCGCAACGTTGTG 88

RESULT 41
US-09-270-767-22149/c
; Sequence 22149, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767

; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6867
; LENGTH: 173
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-22149/c

Query Match      68.0%; Score 17; DB 3; Length 173;
Best Local Similarity 80.0%; Pred. No. 95;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAACTGCATCCCAATGTTATG 25
   ||||| ||||| ||||| ||||| |||||
Db 112 AAACAACTGCAACCGCAACGTTGTG 88

; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 22149
; LENGTH: 173
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-22149

Query Match      68.0%; Score 17; DB 3; Length 173;
Best Local Similarity 80.0%; Pred. No. 95;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAACTGCATCCCAATGTTATG 25
   ||||| ||||| ||||| ||||| |||||
Db 112 AAACAACTGCAACCGCAACGTTGTG 88

RESULT 42
US-09-134-000C-150/c
; Sequence 150, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 150
; LENGTH: 243
; TYPE: DNA
; ORGANISM: Enterococcus faecalis
US-09-134-000C-150

Query Match      68.0%; Score 17; DB 3; Length 243;
Best Local Similarity 80.0%; Pred. No. 1e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAACTGCATCCCAATGTTATG 25
   ||||| ||||| ||||| ||||| |||||
Db 67 AAATATCTGCATCCCAATGATAGG 43

RESULT 43
US-09-270-767-4566
; Sequence 4566, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4566
; LENGTH: 430
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-4566

Query Match      68.0%; Score 17; DB 3; Length 430;
Best Local Similarity 80.0%; Pred. No. 1.2e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAACTGCATCCCAATGTTATG 25
   ||||| ||||| ||||| ||||| |||||
Db 312 AAAGAACTGCATCGCCCAATATCATG 336
```

```

RESULT 44
US-09-270-767-19848
; Sequence 19848, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 19848
; LENGTH: 430
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-19848

Query Match          68.0%; Score 17; DB 3; Length 430;
Best Local Similarity 80.0%; Pred.No. 1.2e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY   1 AAAAACTGCATCCCAATGTTATG 25
      ||| ||||| ||| | |||
DB   312 AAAGAACTGCATGCCAATTATCATG 336

RESULT 45
US-09-949-016-131417/c
; Sequence 131417, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CU001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 131417
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-131417

Query Match          68.0%; Score 17; DB 3; Length 601;
Best Local Similarity 80.0%; Pred.No. 1.3e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY   1 AAAAAACTGCATCCCAATGTTATG 25
      ||| ||||| ||| ||||| |||
DB   514 AAAAACATTAATCCCAATGTTTTG 490

RESULT 46
US-09-543-681A-3477
; Sequence 3477, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543.681A

```

Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAACTGCATCCCAATGTTATG 25
|||||
Db 8077 AAAAAATGCATGCATATGTAATG 8101
|||||

Db 21415 ACAAACCTGCAGCCCCCAAGGAATG 21391
|||||

Search completed: February 3, 2006, 16:32:41
Job time : 91.6667 secs

RESULT 49

US-09-949-016-12528
; Sequence 12528, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12528
; LENGTH: 76810
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(76810)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12528

Query Match 68.0%; Score 17; DB 3; Length 76810;

Best Local Similarity 80.0%; Pred. No. 3.8e+02; Mismatches 5; Indels 0; Gaps 0;

Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAACTGCATCCCAATGTTATG 25
|||||
Db 41677 AAAAACTGCATAACCCATGTGATG 41701
|||||

RESULT 50

US-09-792-616-1/c
; Sequence 1, Application US/09792616
; Patent No. 6780587
; GENERAL INFORMATION:
; APPLICANT: PXE International, Inc.
; APPLICANT: University of Hawaii
; TITLE OF INVENTION: Mutations in a gene encoding an ABC transporter (MRP6) causing
; TITLE OF INVENTION: Pseudoxanthoma Elasticum
; FILE REFERENCE: PXE-001
; CURRENT APPLICATION NUMBER: US/09/792,616
; CURRENT FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 107820
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: "n" can be an A or a T or a G or a C
US-09-792-616-1

Query Match 68.0%; Score 17; DB 3; Length 107820;

Best Local Similarity 80.0%; Pred. No. 4.1e+02; Mismatches 5; Indels 0; Gaps 0;

Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAACTGCATCCCAATGTTATG 25

GenCore version 5.1.7 Copyright (c) 1993 - 2006 Bioceleration Ltd.									
OM nucleic - nucleic search, using sw model									
Run on: February 3, 2006, 14:37:28 ; Search time 365.556 Seconds (without alignments) 565.535 Million cell updates/sec									
Title: US-10-719-900-25									
Perfect score: 25									
Sequence: 1 aaaaactgcatacccaatgttatg 25									
Scoring table: IDENTITY_NUC									
Gapop 10.0 , Gapext 1.0									
Searched: 9793542 seqs, 4134689005 residues									
Total number of hits satisfying chosen parameters: 19587084									
Minimum DB seq length: 0									
Maximum DB seq length: 2000000000									
Post-processing: Minimum Match 0%									
Maximum Match 100%									
Listing first 150 summaries									
Database : Published Applications NA Main:									
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:									
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:									
3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:									
4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:									
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:									
6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:									
7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:									
8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:									
9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:									
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:									
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.									
SUMMARIES									
Result	Query	Score	Match	Length	DB	ID	Description		
1	25	100.0	25	8	US-10-719-900-25	Sequence 25, Appl			
2	24	96.0	25	7	US-10-719-956-283109	Sequence 283109,			
3	23.4	93.6	25	8	US-10-719-900-26	Sequence 26, Appl			
4	22.4	89.6	25	7	US-10-719-956-283110	Sequence 283110,			
5	22.4	89.6	3977	6	US-10-400-435A-1	Sequence 1, Appli			
6	21.8	87.2	270	6	US-10-305-720-1010	Sequence 1010, Ap			
7	20.4	81.6	1914	7	US-10-282-122A-34587	Sequence 34587, A			
8	20.4	81.6	1953	7	US-10-724-972A-1131	Sequence 1131, Ap			
9	20.2	80.8	619	7	US-10-424-599-111654	Sequence 111654,			
10	19.8	79.2	1904	7	US-10-424-599-111652	Sequence 111652,			
11	19.2	76.8	619	4	US-09-925-065A-510687	Sequence 510687,			
12	18.8	75.2	675	4	US-09-925-065A-905792	Sequence 905792,			
13	18.8	75.2	53542	3	US-09-801-574-61	Sequence 61, Appl			
14	18.6	74.4	168174	7	US-10-071-411-63	Sequence 63, Appl			
15	18.6	74.4	168174	8	US-10-914-799-63	Sequence 63, Appl			
16	18.6	74.4	168273	7	US-10-071-411-2	Sequence 2, Appli			
17	18.6	74.4	168273	8	US-10-914-799-2	Sequence 2, Appli			
18	18.2	72.8	288	7	US-10-424-599-2458	Sequence 2458, Ap			
19	18.2	72.8	488	4	US-09-925-065A-603866	Sequence 603866,			
20	18.2	72.8	488	4	US-09-925-065A-603867	Sequence 603867,			
21	18.2	72.8	488	4	US-09-925-065A-603868	Sequence 603868,			
22	18.2	72.8	545	4	US-09-925-065A-133035	Sequence 133035,			
23	18.2	72.8	580	4	US-09-925-065A-788458	Sequence 788458,			

Sequence 367029, A
Sequence 38192, A
Sequence 427422, A
Sequence 6757, Ap
Sequence 160, App
Sequence 5682, Ap
Sequence 17730, A
Sequence 2455, Ap
Sequence 133034, A
Sequence 271268, A
Sequence 271269, A
Sequence 271270, A
Sequence 271271, A
Sequence 18209, A
Sequence 18210, A
Sequence 18209, A
Sequence 18210, A
Sequence 427423, A
Sequence 12160, A
Sequence 668046, A
Sequence 89164, A
Sequence 6890, Ap
Sequence 6234, Ap
Sequence 62193, A
Sequence 26833, A
Sequence 464930, A
Sequence 180366, A
Sequence 180366, A
Sequence 395978, A
Sequence 395978, A
Sequence 895797, A
Sequence 895797, A
Sequence 1132, Ap
Sequence 1132, Ap
Sequence 56255, A
Sequence 364719, A
Sequence 364720, A
Sequence 268, App
Sequence 268, App
Sequence 288, App
Sequence 798669, A
Sequence 230053, A
Sequence 230053, A
Sequence 230053, A
Sequence 54943, A
Sequence 41816, A
Sequence 58232, A
Sequence 30931, A
Sequence 839, App
Sequence 101593, A
Sequence 101593, A
Sequence 68, Appl
Sequence 12, Appl
Sequence 2903, Ap
Sequence 2903, Ap
Sequence 1182, Ap
Sequence 1182, Ap
Sequence 2567, Ap
Sequence 255, App
Sequence 6713, Ap
Sequence 3, Appli
Sequence 17745, A
Sequence 10, Appl
Sequence 2058, Ap
Sequence 1, Appli
Sequence 354661, A
Sequence 78, Appl
Sequence 78, Appl
Sequence 19906, A
Sequence 574, App
Sequence 31318, A

```

97 17.2 68.8 311 8 US-10-674-124A-26606
98 17.2 68.8 427 8 US-10-767-795-1414
c 99 17.2 68.8 451 7 US-10-437-963-91182
100 17.2 68.8 495 8 US-10-425-115-53972
c 101 17.2 68.8 513 7 US-10-437-963-52090
102 17.2 68.8 570 4 US-09-925-065A-880166
103 17.2 68.8 620 4 US-09-925-065A-341098
104 17.2 68.8 624 4 US-09-925-065A-341099
c 105 17.2 68.8 624 4 US-09-925-065A-736692
106 17.2 68.8 634 4 US-09-925-065A-754089
c 107 17.2 68.8 634 4 US-09-925-065A-745337
108 17.2 68.8 645 5 US-10-027-632-240922
109 17.2 68.8 645 5 US-10-027-632-240923
110 17.2 68.8 645 5 US-10-027-632-240924
111 17.2 68.8 645 5 US-10-027-632-240925
112 17.2 68.8 645 5 US-10-027-632-240922
113 17.2 68.8 645 6 US-10-027-632-240923
114 17.2 68.8 645 6 US-10-027-632-240924
115 17.2 68.8 645 6 US-10-027-632-240925
c 116 17.2 68.8 681 7 US-10-282-122A-10935
c 117 17.2 68.8 975 5 US-10-027-632-116833
c 118 17.2 68.8 975 5 US-10-027-632-116833
c 119 17.2 68.8 1012 9 US-10-499-184-11
c 120 17.2 68.8 1039 5 US-10-027-632-121527
c 121 17.2 68.8 1039 5 US-10-027-632-121528
c 122 17.2 68.8 1039 5 US-10-027-632-121529
c 123 17.2 68.8 1039 6 US-10-027-632-121527
c 124 17.2 68.8 1039 6 US-10-027-632-121528
c 125 17.2 68.8 1039 6 US-10-027-632-121529
c 126 17.2 68.8 1875 7 US-10-260-238-903
c 127 17.2 68.8 7422 10 US-11-097-143-10606
c 128 17.2 68.8 7447 10 US-11-097-143-25978
c 129 17.2 68.8 8132 10 US-11-097-143-25981
c 130 17.2 68.8 36714 7 US-10-322-281-864
c 131 17.2 68.8 38605 5 US-10-087-192-304
c 132 17.2 68.8 62231 3 US-10-087-192-493
c 133 17.2 68.8 96588 3 US-09-997-722-292
c 134 17.2 68.8 268685 6 US-10-265-071-22
c 135 17.2 68.8 268685 6 US-10-025-966A-22
c 136 17.2 68.8 775062 8 US-10-719-993-6844
c 137 17 68.0 201 8 US-10-719-993-8989
c 138 17 68.0 201 8 US-10-719-993-8991
c 139 17 68.0 201 8 US-10-719-993-8992
c 140 17 68.0 201 8 US-10-719-993-8993
c 141 17 68.0 409 4 US-09-925-065A-172463
c 142 17 68.0 410 4 US-09-925-065A-592812
c 143 17 68.0 426 5 US-10-027-632-182442
c 144 17 68.0 426 5 US-10-027-632-182442
c 145 17 68.0 426 6 US-10-027-632-182442
c 146 17 68.0 426 6 US-10-027-632-307207
c 147 17 68.0 427 5 US-10-027-632-79765
c 148 17 68.0 427 5 US-10-027-632-109526
c 149 17 68.0 427 6 US-10-027-632-79765
c 150 17 68.0 427 6 US-10-027-632-109526

Sequence 26606, A
Sequence 1414, Ap
Sequence 91182, A
Sequence 53972, A
Sequence 52090, A
Sequence 880166,
Sequence 341098,
Sequence 341099,
Sequence 736692,
Sequence 754089,
Sequence 745337,
Sequence 240922,
Sequence 240923,
Sequence 240924,
Sequence 240925,
Sequence 10935, A
Sequence 116833,
Sequence 116833,
Sequence 11, Appl
Sequence 121527,
Sequence 121528,
Sequence 121529,
Sequence 121527,
Sequence 121528,
Sequence 121529,
Sequence 903, App
Sequence 10606, A
Sequence 25978, A
Sequence 25981, A
Sequence 864, App
Sequence 304, App
Sequence 493, App
Sequence 292, App
Sequence 22, Appl
Sequence 22, Appl
Sequence 6844, App
Sequence 8989, App
Sequence 8991, App
Sequence 8992, App
Sequence 8993, App
Sequence 172463,
Sequence 592812,
Sequence 182442,
Sequence 307207,
Sequence 182442,
Sequence 307207,
Sequence 79765, A
Sequence 109526,
Sequence 79765, A
Sequence 109526,

Query Match 100.0%; Score 25; DB 8; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.17; 0; Indels 0; Gaps 0;
Matches 25; Conservative 0; Mismatches 0;

QY 1 AAAAACTGCATCCCAATGTTATG 25
Db 1 AAAAACTGCATCCCAATGTTATG 25

RESULT 2
US-10-719-956-283109
; Sequence 283109, Application US/10719956
; Publication No. US20040146910A1
; GENERAL INFORMATION:
; APPLICANT: Xue Mei Zhou
; TITLE OF INVENTION: Methods of Genetic Analysis of Rat
; FILE REFERENCE: 3527.1
; CURRENT APPLICATION NUMBER: US/10/719,956
; CURRENT FILING DATE: 2003-11-20
; PRIOR APPLICATION NUMBER: 60/427,836
; PRIOR FILING DATE: 2002 11 20
; NUMBER OF SEQ ID NOS: 699466
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 283109
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-10-719-956-283109

Query Match 96.0%; Score 24; DB 7; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.49;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AAAAACTGCATCCCAATGTTATG 25
Db 2 AAAAACTGCATCCCAATGTTATG 25

RESULT 3
US-10-719-900-25
; Sequence 25, Application US/10719900
; Publication No. US20050026164A1
; GENERAL INFORMATION:
; APPLICANT: Xue Mei Zhou
; TITLE OF INVENTION: Methods of Genetic Analysis of Mouse
; FILE REFERENCE: 3528.1
; CURRENT APPLICATION NUMBER: US/10/719,900
; CURRENT FILING DATE: 2003-11-20
; PRIOR APPLICATION NUMBER: 60/427,808
; PRIOR FILING DATE: 2002 11 20
; NUMBER OF SEQ ID NOS: 982914
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 26
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-719-900-26

Query Match 93.6%; Score 23.4; DB 8; Length 25;
Best Local Similarity 96.0%; Pred. No. 0.93;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAAACTGCATCCCAATGTTATG 25
Db 1 AAAAACTGCATCCCAATGTTATG 25
```

ALIGNMENTS

```

RESULT 1
US-10-719-900-25
; Sequence 25, Application US/10719900
; Publication No. US20050026164A1
; GENERAL INFORMATION:
; APPLICANT: Xue Mei Zhou
; TITLE OF INVENTION: Methods of Genetic Analysis of Mouse
; FILE REFERENCE: 3528.1
; CURRENT APPLICATION NUMBER: US/10/719,900
; CURRENT FILING DATE: 2003-11-20
; PRIOR APPLICATION NUMBER: 60/427,808
; PRIOR FILING DATE: 2002 11 20
; NUMBER OF SEQ ID NOS: 982914
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
```


RESULT 4

US-10-719-956-283110
; Sequence 283110, Application US/10719956
; Publication No. US20040146910A1
; GENERAL INFORMATION:
; APPLICANT: Xue Mei Zhou
; TITLE OF INVENTION: Methods of Genetic Analysis of Rat
; FILE REFERENCE: 3527.1
; CURRENT APPLICATION NUMBER: US/10/719,956
; CURRENT FILING DATE: 2003-11-20
; PRIOR APPLICATION NUMBER: 60/427,836
; PRIOR FILING DATE: 2002 11 20
; NUMBER OF SEQ ID NOS: 699466
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 283110
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-10-719-956-283110

Query Match 89.6%; Score 22.4; DB 7; Length 25;
Best Local Similarity 95.8%; Pred. No. 2.7;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 AAAAAGTCATCCCAATGTTATG 25
|||||
Db 2 AAAAAGTCATCCCAATGTTATG 25

RESULT 5

US-10-400-435A-1
; Sequence 1, Application US/10400435A
; Publication No. US20030224416A1
; GENERAL INFORMATION:
; APPLICANT: Cochran, Susan
; APPLICANT: Yamagami, Keiji
; APPLICANT: Ohashi, Yoshitaka
; TITLE OF INVENTION: SCHIZOPHRENIA RELATED GENE
; FILE REFERENCE: 2003-0437A/WNC/00279
; CURRENT APPLICATION NUMBER: US/10/400,435A
; CURRENT FILING DATE: 2003-03-28
; PRIOR APPLICATION NUMBER: EP02007114.8
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 1
; LENGTH: 3977
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1162)..(2919)
; FEATURE:
; NAME/KEY: gene
; LOCATION: (1)..(3977)
; FEATURE:
; NAME/KEY: polyA signal
; LOCATION: (3958)..(3963)
; PUBLICATION INFORMATION:
; AUTHORS: Lunau C.J., Williams, J.B., Marshall, J., Levitan, E.S., Oliva, C.,
; AUTHORS: Smith, J.S., Antanavage, J., Folander, K., Stein, R.B., Swanson, R., Kaczmarek, L.
; TITLE: Alternate splicing contributes to K+ channel diversity in the mammalian
; TITLE: central nervous system.
; JOURNAL: Proc. Natl. Acad. Sci. USA.
; VOLUME: 88
; ISSUE: 5
; PAGES: 3932-3936
; DATABASE ACCESSION NUMBER: NM_012856
; DATABASE ENTRY DATE: 1991-05-01
US-10-400-435A-1

Query Match 89.6%; Score 22.4; DB 6; Length 3977;

Best Local Similarity 95.8%; Pred. No. 6.7;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 AAAAAGTCATCCCAATGTTATG 25
|||||
Db 3915 AAAAAGTCATCCCAATGTTATG 3938

RESULT 6

US-10-305-720-1010
; Sequence 1010, Application US/10305720
; Publication No. US20040010136A1
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice K.; Seilhamer, Jeffrey J.
; TITLE OF INVENTION: Composition for the Detection of Signaling Pathway Gene Expression
; FILE REFERENCE: PA-0002-1 CON
; CURRENT APPLICATION NUMBER: US/10/305,720
; CURRENT FILING DATE: 2002-11-26
; PRIOR APPLICATION NUMBER: 09/016,434
; PRIOR FILING DATE: 1998-01-30
; NUMBER OF SEQ ID NOS: 1490
; SOFTWARE: PERL Program
; SEQ ID NO 1010
; LENGTH: 270
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20040010136A1 926642
US-10-305-720-1010

Query Match 87.2%; Score 21.8; DB 6; Length 270;
Best Local Similarity 92.0%; Pred. No. 7.8;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AAAAAGTCATCCCAATGTTATG 25
|||||
Db 164 AAAAAGTCATCCCAATGTTATG 188

RESULT 7

US-10-282-122A-34587
; Sequence 34587, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23

```
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 34587
; LENGTH: 1914
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-10-282-122A-34587

Query Match      81.6%; Score 20.4; DB 7; Length 1914;
Best Local Similarity 95.5%; Pred. No. 49;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 AAAACTGCATCCCAATGTTAT 24
|||||
Db 1441 AAAACTGCATCCTCAATGTTAT 1462

RESULT 8
US-10-724-972A-1131
; Sequence 1131, Application US/10724972A
; Publication No. US20040147734A1
; GENERAL INFORMATION:
; APPLICANT: Doucette-Stamm, Lynn
; APPLICANT: Bush, David
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/10/724,972A
; CURRENT FILING DATE: 2003-12-01
; PRIOR APPLICATION NUMBER: 09/450,969
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: 09/134,001
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 7544
; SEQ ID NO 1131
; LENGTH: 1953
; TYPE: DNA
; ORGANISM: S.epidermidis
US-10-724-972A-1131

Query Match      81.6%; Score 20.4; DB 7; Length 1953;
Best Local Similarity 95.5%; Pred. No. 49;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 AAAACTGCATCCCAATGTTAT 24
|||||
Db 1477 AAAACTGCATCCTCAATGTTAT 1498

RESULT 9
US-10-424-599-111654
; Sequence 111654, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: Plants and Uses Thereof for Plant Improvement
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
```

```
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 111654
; LENGTH: 619
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_71833C.1
US-10-424-599-111654

Query Match      80.8%; Score 20.2; DB 7; Length 619;
Best Local Similarity 88.0%; Pred. No. 49;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAACTGCATCCCAATGTTATG 25
|||||
Db 45 ATAAAAATGCATCCCAATGTTAAG 69

RESULT 10
US-10-424-599-111652
; Sequence 111652, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 111652
; LENGTH: 1904
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_71831C.1
US-10-424-599-111652

Query Match      79.2%; Score 19.8; DB 7; Length 1904;
Best Local Similarity 91.3%; Pred. No. 92;
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAACTGCATCCCAATGTTA 23
|||||
Db 1292 ATAAAAATGCATCCCAATGTTA 1314

RESULT 11
US-09-925-065A-510687
; Sequence 510687, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: Nucleotide Polymorphisms in the Human Genome
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
```

; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 510687
; LENGTH: 619
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-510687

Query Match 76.8%; Score 19.2; DB 4; Length 619;
Best Local Similarity 87.5%; Pred. No. 1.4e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAACTGCATCCCAATGTTAT 24
DB 142 AAAAAATATGATCCCAATGTTAT 165

RESULT 12
US-09-925-065A-905792
; Sequence 905792, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 905792
; LENGTH: 675
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-905792

Query Match 75.2%; Score 18.8; DB 4; Length 675;
Best Local Similarity 90.9%; Pred. No. 2.2e+02;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAACTGCATCCCAATGTT 22
DB 530 AATAAACTGCATCCCAATTT 551

RESULT 13
US-09-801-574-61
; Sequence 61, Application US/09801574
; Patent No. US20020081592A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Peijiang Jeremy
; TITLE OF INVENTION: Reproduction-Specific Genes
; FILE REFERENCE: 0399.2007-002
; CURRENT APPLICATION NUMBER: US/09/801,574
; CURRENT FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: 60/187,518
; PRIOR FILING DATE: 2000-03-07
; PRIOR APPLICATION NUMBER: 60/261,557
; PRIOR FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 61
; LENGTH: 53542
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-801-574-61

Query Match 75.2%; Score 18.8; DB 3; Length 53542;
Best Local Similarity 90.9%; Pred. No. 4.8e+02;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 AAATGTCATCCCAATGTTATG 25
DB 29516 AAATTCATCCCAATGATG 29537

RESULT 14
US-10-071-411-63
; Sequence 63, Application US/10071411
; Publication No. US20040170974A1
; GENERAL INFORMATION:
; APPLICANT: Joanne Meyer
; TITLE OF INVENTION: Detection of Polymorphisms in the Human
; FILE REFERENCE: MRI-021
; CURRENT APPLICATION NUMBER: US/10/071,411
; CURRENT FILING DATE: 2002-02-07
; PRIOR APPLICATION NUMBER: 60/267,515
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: 60/314,248
; PRIOR FILING DATE: 2001-08-21
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 63
; LENGTH: 168174
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(168174)
; OTHER INFORMATION: n = A,T,C or G
US-10-071-411-63

Query Match 74.4%; Score 18.6; DB 7; Length 168174;
Best Local Similarity 84.0%; Pred. No. 7.3e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAACTGCATCCCAATGTTATG 25
DB 91147 AAAGAACTGCATCCTTAATGTTCTG 91171

RESULT 15
US-10-914-799-63
; Sequence 63, Application US/10914799
; Publication No. US20050009084A1
; GENERAL INFORMATION:
; APPLICANT: Joanne Meyer
; TITLE OF INVENTION: Detection of Polymorphisms in the Human
; FILE REFERENCE: MRI-021
; CURRENT APPLICATION NUMBER: US/10/914,799
; CURRENT FILING DATE: 2004-08-09
; PRIOR APPLICATION NUMBER: US/10/071,411
; PRIOR FILING DATE: 2002-02-07
; PRIOR APPLICATION NUMBER: 60/267,515
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: 60/314,248
; PRIOR FILING DATE: 2001-08-21
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 63


```
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 603866
; LENGTH: 488
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-603866

Query Match          72.8%; Score 18.2; DB 4; Length 488;
Best Local Similarity 87.0%; Pred. No. 3.9e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AAAAACTGCATCCCAATGTTA 23
    ||||| ||||| ||||| |||||
Db 40 AAAAAAAGCATGCCCAATGTTA 62

RESULT 20
US-09-925-065A-603867
; Sequence 603867, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 603867
; LENGTH: 488
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-603867

Query Match          72.8%; Score 18.2; DB 4; Length 488;
Best Local Similarity 87.0%; Pred. No. 3.9e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AAAAACTGCATCCCAATGTTA 23
    ||||| ||||| ||||| |||||
Db 40 AAAAAAAGCATGCCCAATGTTA 62

RESULT 21
US-09-925-065A-603868
; Sequence 603868, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30

; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 603867
; LENGTH: 488
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-603868

Query Match          72.8%; Score 18.2; DB 4; Length 488;
Best Local Similarity 87.0%; Pred. No. 3.9e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AAAAACTGCATCCCAATGTTA 23
    ||||| ||||| ||||| |||||
Db 40 AAAAAAAGCATGCCCAATGTTA 62

RESULT 22
US-09-925-065A-133035/c
; Sequence 133035, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 133035
; LENGTH: 545
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-133035

Query Match          72.8%; Score 18.2; DB 4; Length 545;
Best Local Similarity 87.0%; Pred. No. 4e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 AAAAACTGCATCCCAATGTTAT 24
    ||||| ||||| ||||| |||||
Db 197 AAAAACTGCATCCCAAAAGTAAT 175

RESULT 23
US-09-925-065A-788458
; Sequence 788458, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
```

```
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; PRIOR APPLICATION NUMBER: US 60/261,766
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 788458
; LENGTH: 580
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-788458
```

```
Query Match 72.8%; Score 18.2; DB 4; Length 580;
Best Local Similarity 87.0%; Pred. No. 4.1e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 1 AAAAAAGTCATCCCAATGTTA 23
    ||||| ||||| ||||| |||||
Db 321 AAAAAAGCATGCCCAATGTTA 343
```

```
RESULT 24
US-09-925-065A-367029/c
; Sequence 367029, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 367029
; LENGTH: 588
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-367029
```

```
Query Match 72.8%; Score 18.2; DB 4; Length 588;
Best Local Similarity 87.0%; Pred. No. 4.1e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 1 AAAAAAGTCATCCCAATGTTA 23
    ||||| ||||| ||||| |||||
Db 301 AAAAAAGCATACACAAAGTTA 279
```

```
RESULT 25
US-09-925-065A-38192
; Sequence 38192, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT FILING DATE: 2001-08-08
```

```
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 38192
; LENGTH: 618
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-38192
```

```
Query Match 72.8%; Score 18.2; DB 4; Length 618;
Best Local Similarity 87.0%; Pred. No. 4.1e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 1 AAAAAAGTCATCCCAATGTTA 23
    ||||| ||||| ||||| |||||
Db 353 AAAAAAGCATGCCCAATGTTA 375
```

```
RESULT 26
US-09-925-065A-427422/c
; Sequence 427422, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 427422
; LENGTH: 633
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-427422
```

```
Query Match 72.8%; Score 18.2; DB 4; Length 633;
Best Local Similarity 87.0%; Pred. No. 4.1e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 1 AAAAAAGTCATCCCAATGTTA 23
    ||||| ||||| ||||| |||||
Db 299 AAAAAAGCATGCCCAATGTTA 277
```

```
RESULT 27
US-10-719-993-6757/c
; Sequence 6757, Application US/10719993
; Publication No. US20040265849A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; ALZHEIMER'S DISEASE, METHODS OF DETECTION AND USES THEREOF
```

; FILE REFERENCE: CL001496
; CURRENT APPLICATION NUMBER: US/10/719,993
; CURRENT FILING DATE: 2003-11-24
; NUMBER OF SEQ ID NOS: 5342
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6757
; LENGTH: 13080
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-719-993-6757

Query Match 72.8%; Score 18.2; DB 8; Length 13080;
Best Local Similarity 87.0%; Pred. No. 7.1e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAACTGCATCCCAATGTTA 23
||||| ||||| ||||| ||||| |||||

DB 3685 AAAACACTGCTTCCCAATGTTA 3663

RESULT 28

US-10-087-192-160
; Sequence 160, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:

; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 160
; LENGTH: 95683
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-087-192-160

Query Match 72.8%; Score 18.2; DB 5; Length 95683;
Best Local Similarity 87.0%; Pred. No. 1e+03;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAACTGCATCCCAATGTTAT 24
||||| ||||| ||||| ||||| |||||

DB 8156 AAATACTGCAACCCCTATGTTAT 8178

RESULT 29

US-10-741-601-5682/c
; Sequence 5682, Application US/10741601
; Publication No. US20040166519A1
; GENERAL INFORMATION:

; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001500
; CURRENT APPLICATION NUMBER: US/10/741,601
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 26415
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5682
; LENGTH: 561515
; TYPE: DNA
; ORGANISM: Homo sapiens

; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(561515)

; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1
US-10-741-601-5682

Query Match 72.8%; Score 18.2; DB 7; Length 561515;
Best Local Similarity 87.0%; Pred. No. 1.3e+03;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAACTGCATCCCAATGTTA 23
||||| ||||| ||||| ||||| |||||

DB 253561 AAAAACCTCTCTCCCAATGTTA 253539

RESULT 30

US-10-741-600-17730/c
; Sequence 17730, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:

; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17730
; LENGTH: 561515
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(561515)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1
US-10-741-600-17730

Query Match 72.8%; Score 18.2; DB 8; Length 561515;
Best Local Similarity 87.0%; Pred. No. 1.3e+03;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAACTGCATCCCAATGTTA 23
||||| ||||| ||||| ||||| |||||

DB 253561 AAAAACCTCTCTCCCAATGTTA 253539

RESULT 31

US-10-424-599-2455
; Sequence 2455, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:

; APPLICANT: Kovalic David J
; APPLICANT: La Rosa Thomas J
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 2455
; LENGTH: 467
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_102222C.1
US-10-424-599-2455

Query Match 71.2%; Score 17.8; DB 7; Length 467;
Best Local Similarity 90.5%; Pred. No. 6e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 AAAACTGCATCCCAATGTTAT 24
||||| ||||| ||||| ||||| |||||

```
Db      399 AAAGTGCATCCCAATTTT 419

RESULT 32
US-09-925-065A-133034/c
; Sequence 133034, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 10827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 133034
; LENGTH: 545
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-133034

Query Match      71.2%; Score 17.8; DB 4; Length 545;
Best Local Similarity 82.6%; Pred. No. 6.1e+02;
Matches 19; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY      2 AAAGTGCATCCCAATGTT 24
Db      197 AAAGTGCATCCCAAGTAA 175

RESULT 33
US-09-925-065A-271268/c
; Sequence 271268, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 10827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-05-09
; PRIOR APPLICATION NUMBER: US 60/289,846
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 271268
; LENGTH: 568
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-271268

Query Match      71.2%; Score 17.8; DB 4; Length 568;
Best Local Similarity 90.5%; Pred. No. 6.2e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      2 AAAGTGCATCCCAATGTT 22
Db      259 AAAGTGCATCCCAATGTT 239

RESULT 34
US-09-925-065A-271269/c
; Sequence 271269, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 10827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-05-09
; PRIOR APPLICATION NUMBER: US 60/289,846
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 271269
; LENGTH: 568
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-271269

Query Match      71.2%; Score 17.8; DB 4; Length 568;
Best Local Similarity 90.5%; Pred. No. 6.2e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      2 AAAGTGCATCCCAATGTT 22
Db      259 AAAGTGCATCCCAATGTT 239

RESULT 35
US-09-925-065A-271270/c
; Sequence 271270, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 10827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 271270
; LENGTH: 568
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-271270
```



```
Query Match          71.2%; Score 17.8; DB 4; Length 568;
Best Local Similarity 90.5%; Pred. No. 6.2e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AAAAACTGCATCCCAATGT 22
   ||||| ||||| |||||
Db 259 AAAAACTCTCCCAATGT 239

RESULT 36
US-09-925-065A-271271/c
; Sequence 271271, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 271271
; LENGTH: 568
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-271271

Query Match          71.2%; Score 17.8; DB 4; Length 568;
Best Local Similarity 90.5%; Pred. No. 6.2e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AAAAACTGCATCCCAATGT 22
   ||||| ||||| |||||
Db 259 AAAAACTCTCCCAATGT 239

RESULT 37
US-10-363-345A-18209/c
; Sequence 18209, Application US/10363345A
; Publication No. US20040234960A1
; GENERAL INFORMATION:
; APPLICANT: Alexander Olek
; APPLICANT: Christian Piepenbrock
; APPLICANT: Kurt Berlin
; TITLE OF INVENTION: Method for determining the degree of methylation of defined
; FILE OF INVENTION: cytosines in genomic DNA in the sequence context of 5'-CpG-3
; FILE REFERENCE: E01/1227
; CURRENT APPLICATION NUMBER: US/10/363,345A
; CURRENT FILING DATE: 2003-03-03
; NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 18209
; LENGTH: 578
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; OTHER INFORMATION: CpG-island No: 18209
US-10-363-345A-18209

Query Match          71.2%; Score 17.8; DB 8; Length 578;
Best Local Similarity 90.5%; Pred. No. 6.2e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 1 AAAAAACTGCATCCCAATGT 21
   ||||| ||||| |||||
Db 403 AAAAAACCGCATCCCAATTT 383

RESULT 38
US-10-363-345A-18210
; Sequence 18210, Application US/10363345A
; Publication No. US20040234960A1
; GENERAL INFORMATION:
; APPLICANT: Alexander Olek
; APPLICANT: Christian Piepenbrock
; APPLICANT: Kurt Berlin
; TITLE OF INVENTION: Method for determining the degree of methylation of defined
; FILE OF INVENTION: cytosines in genomic DNA in the sequence context of 5'-CpG-3
; FILE REFERENCE: E01/1227
; CURRENT APPLICATION NUMBER: US/10/363,345A
; CURRENT FILING DATE: 2003-03-03
; NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 18210
; LENGTH: 578
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; OTHER INFORMATION: CpG-island No: 18210
US-10-363-345A-18210

Query Match          71.2%; Score 17.8; DB 8; Length 578;
Best Local Similarity 90.5%; Pred. No. 6.2e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAACTGCATCCCAATGT 21
   ||||| ||||| |||||
Db 176 AAAAAACCGCATCCCAATTT 196

RESULT 39
US-10-363-483A-18209/c
; Sequence 18209, Application US/10363483A
; Publication No. US20050064401A1
; GENERAL INFORMATION:
; APPLICANT: Alexander Olek
; APPLICANT: Christian Piepenbrock
; APPLICANT: Kurt Berlin
; TITLE OF INVENTION: Diagnosis of illnesses or predisposition to certain
; FILE OF INVENTION: illnesses
; FILE REFERENCE: 82011
; CURRENT APPLICATION NUMBER: US/10/363,483A
; CURRENT FILING DATE: 2003-03-03
; NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 18209
; LENGTH: 578
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; OTHER INFORMATION: CpG-island No: 18209
US-10-363-483A-18209

Query Match          71.2%; Score 17.8; DB 9; Length 578;
Best Local Similarity 90.5%; Pred. No. 6.2e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAACTGCATCCCAATGT 21
   ||||| ||||| |||||
Db 403 AAAAAACCGCATCCCAATTT 383

RESULT 40
US-10-363-483A-18210
; Sequence 18210, Application US/10363483A
```

```
; Publication No. US20050064401A1
; GENERAL INFORMATION:
; APPLICANT: Alexander Olek
; APPLICANT: Christian Piepenbrock
; APPLICANT: Kurt Berlin
; TITLE OF INVENTION: Diagnosis of illnesses or predisposition to certain
; TITLE OF INVENTION: illnesses
; FILE REFERENCE: 82011
; CURRENT APPLICATION NUMBER: US/10/363,483A
; CURRENT FILING DATE: 2003-03-03
; NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 18210
; LENGTH: 578
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; OTHER INFORMATION: CpG-island No: 18210
US-10-363-483A-18210

Query Match          71.2%; Score 17.8; DB 9; Length 578;
Best Local Similarity 90.5%; Pred. No. 6.2e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAACTGCATCCCAATGT 21
   ||||| ||||| ||||| |||||
Db 176 AAAAAACCGCATCCCAATTT 196

RESULT 41
US-09-925-065A-427423/c
; Sequence 427423, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 427423
; LENGTH: 633
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-427423

Query Match          71.2%; Score 17.8; DB 4; Length 633;
Best Local Similarity 82.6%; Pred. No. 6.3e+02;
Matches 19; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAACTGCATCCCAATGTTA 23
   ||||| ||||| ||||| |||||
Db 299 AAAAAAAGCATGCCCAATGTTA 277

RESULT 42
US-10-282-122A-12160/c
; Sequence 12160, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
```

```
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12160
; LENGTH: 1083
; TYPE: DNA
; ORGANISM: Bacteroides fragilis
US-10-282-122A-12160

Query Match          71.2%; Score 17.8; DB 7; Length 1083;
Best Local Similarity 90.5%; Pred. No. 6.9e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AAAAACTGCATCCCAATGTT 22
   ||||| ||||| ||||| |||||
Db 633 AAAAACTGCATCCCAATCTT 613

RESULT 43
US-09-925-065A-668046/c
; Sequence 668046, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
```

; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 668046
; LENGTH: 3517
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-668046

Query Match 71.2%; Score 17.8; DB 4; Length 3517;
Best Local Similarity 90.5%; Pred. No. 8.6e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 AAATGATCCCAATGTTAT 24
DB 505 AAATGATCCCAATGTTAT 485

RESULT 44
US-10-437-963-89164
; Sequence 89164, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Brad
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 89164
; LENGTH: 4283
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_87946C.1
US-10-437-963-89164

Query Match 71.2%; Score 17.8; DB 7; Length 4283;
Best Local Similarity 90.5%; Pred. No. 8.9e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 AACTGCATCCCAATGTTATG 25
DB 973 AAATGATCCCAATGTTAAG 993

RESULT 45
US-10-719-993-6890
; Sequence 6890, Application US/10719993
; Publication No. US20040265849A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: ALZHEIMER'S DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001496
; CURRENT APPLICATION NUMBER: US/10/719,993
; CURRENT FILING DATE: 2003-11-24
; NUMBER OF SEQ ID NOS: 55342
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6890
; LENGTH: 303172
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature

; LOCATION: (1)...(303172)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1
US-10-719-993-6890

Query Match 71.2%; Score 17.8; DB 8; Length 303172;
Best Local Similarity 90.5%; Pred. No. 1.8e+03;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAAAGTCATCCCAATGT 21
DB 127063 AAAAAAAGTCATCCCAATGT 127083

RESULT 46
US-10-424-599-6234
; Sequence 6234, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 6234
; LENGTH: 379
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)...(379)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_105639C.1
US-10-424-599-6234

Query Match 70.4%; Score 17.6; DB 7; Length 379;
Best Local Similarity 83.3%; Pred. No. 7.1e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AAAAACTGCATCCCAATGTTATG 25
DB 109 AAGACTTCATCCCAATCTTACG 132

RESULT 47
US-10-425-115-62193/c
; Sequence 62193, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 62193
; LENGTH: 389
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_156715C.1
US-10-425-115-62193

Query Match 70.4%; Score 17.6; DB 8; Length 389;

Best Local Similarity 83.3%; Pred. No. 7.1e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 4; Gaps 0;
QY 2 AAAAAGTCATCCCAATGTTATG 25
DB 372 AAACACTGCATCTCTATGTAATG 349

RESULT 48
US-09-918-995-26833/c
; Sequence 26833, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hvaeq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 26833
; LENGTH: 461
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(461)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-26833

Query Match 70.4%; Score 17.6; DB 3; Length 461;
Best Local Similarity 83.3%; Pred. No. 7.3e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 2 AAAAAGTCATCCCAATGTTATG 25
DB 209 AAAAAGTCATCAACACTGTTAAG 186

RESULT 49
US-09-925-065A-464930/c
; Sequence 464930, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 464930
; LENGTH: 519
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-464930

Query Match 70.4%; Score 17.6; DB 4; Length 519;
Best Local Similarity 83.3%; Pred. No. 7.5e+02;

Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 AAAAAGTCATCCCAATGTTAT 24
DB 503 AAAATGTCATCCCAATGTTT 480

RESULT 50
US-10-027-632-180366/c
; Sequence 180366, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 180366
; LENGTH: 520
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-180366

Query Match 70.4%; Score 17.6; DB 5; Length 520;
Best Local Similarity 83.3%; Pred. No. 7.5e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 AAAAAGTCATCCCAATGTTAT 24
DB 268 AAAATACTGGATCCCAAGATAT 245

Search completed: February 3, 2006, 15:44:02
Job time : 375.556 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 3, 2006, 11:35:31 ; Search time 343.111 Seconds
(without alignments)
61.064 Million cell updates/sec

Title: US-10-719-900-25

Perfect score: 25

Sequence: 1 aaaaactgcattcccaatgttatg 25

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 6068529 seqs, 419036697 residues

Total number of hits satisfying chosen parameters: 12137058

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database : Published Applications NA.New.*

- 1: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
- 2: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
- 5: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
- 7: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
- 9: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
- 10: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
- 11: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	22.4	89.6	1400	8	US-11-136-527-6248
2	22.4	89.6	4019	8	US-11-136-527-2152
3	20.4	81.6	1917	7	US-10-793-626-1467
4	20.4	81.6	3276	7	US-10-793-626-4240
5	19.8	79.2	95050	7	US-10-857-780-7
6	19.4	77.6	25	8	US-11-136-527-295557
7	17.6	70.4	201	7	US-10-995-561-55567
8	17.6	70.4	709	7	US-10-750-185-36949
9	17.6	70.4	709	7	US-10-750-623-36949
10	17.6	70.4	720	7	US-10-750-185-48171
11	17.6	70.4	720	7	US-10-750-623-48171
12	17.2	68.8	201	8	US-10-995-561-55566
13	17.2	68.8	201	8	US-11-124-368A-19759
14	17.2	68.8	201	8	US-11-124-368A-19777
15	17.2	68.8	600	7	US-10-750-185-1670
16	17.2	68.8	600	7	US-10-750-623-1670
17	17.2	68.8	855	7	US-10-750-185-41308
18	17.2	68.8	855	7	US-10-750-623-41308
19	17.2	68.8	1130	7	US-10-750-185-52076
20	17.2	68.8	1130	7	US-10-750-623-52076
21	17.2	68.8	79134	8	US-11-124-368A-2924
22	17.2	68.8	218821	8	US-11-121-086-31

Sequence 13327, A	7	US-10-995-561-13327	68.8	244196	17.2	C 23
Sequence 22, Appl	7	US-10-933-025-22	68.8	268685	17.2	C 24
Sequence 827, App	7	US-10-750-185-827	68.0	600	17	C 25
Sequence 827, App	7	US-10-750-623-827	68.0	600	17	C 26
Sequence 73, Appl	8	US-11-121-086-73	68.0	152335	17	C 27
Sequence 13233, A	7	US-10-995-561-13233	68.0	199130	17	C 28
Sequence 3, Appl1	8	US-11-102-978-3	68.0	340000	17	C 29
Sequence 64, Appl	8	US-11-121-086-64	67.2	142605	16.8	C 30
Sequence 13215, A	8	US-11-121-086-63	67.2	169725	16.8	C 31
Sequence 19474, A	7	US-10-995-561-13215	67.2	241805	16.8	C 32
Sequence 19474, A	7	US-10-750-185-19474	66.4	600	16.6	C 33
Sequence 24537, A	7	US-10-750-623-24537	66.4	600	16.6	C 34
Sequence 24537, A	7	US-10-750-185-24537	66.4	1463	16.6	C 35
Sequence 94, Appl	10	US-11-033-764-94	66.4	2161	16.6	C 36
Sequence 29455, A	7	US-10-750-185-29455	66.4	2524	16.6	C 37
Sequence 29455, A	7	US-10-750-623-29455	66.4	2524	16.6	C 38
Sequence 5, Appl1	8	US-11-192-967-5	66.4	5592	16.6	C 39
Sequence 53, Appl1	8	US-11-193-715-5	66.4	5592	16.6	C 40
Sequence 61, Appl	8	US-11-112-908-63	66.4	149111	16.6	C 41
Sequence 61, Appl	8	US-11-112-908-61	66.4	159497	16.6	C 42
Sequence 44, Appl	8	US-11-121-086-44	66.4	167116	16.6	C 43
Sequence 43, Appl	8	US-11-121-086-43	66.4	175416	16.6	C 44
Sequence 20, Appl	8	US-11-112-908-20	66.4	191331	16.6	C 45
Sequence 56163, A	7	US-10-750-185-56163	65.6	1239	16.4	C 46
Sequence 56163, A	7	US-10-750-623-56163	65.6	1239	16.4	C 47
Sequence 2599, Ap	8	US-11-136-527-2599	65.6	2217	16.4	C 48
Sequence 39, Appl	8	US-11-121-086-39	65.6	207835	16.4	C 49
Sequence 40, Appl	8	US-11-121-086-40	65.6	207835	16.4	C 50
Sequence 3061, Ap	7	US-10-750-185-3061	64.8	600	16.2	C 51
Sequence 3061, Ap	7	US-10-750-623-3061	64.8	600	16.2	C 52
Sequence 3367, Ap	8	US-11-128-061-3367	64.8	978	16.2	C 53
Sequence 7009, Ap	8	US-11-128-061-7009	64.8	978	16.2	C 54
Sequence 3367, Ap	8	US-11-128-049-3367	64.8	978	16.2	C 55
Sequence 7009, Ap	8	US-11-128-049-7009	64.8	978	16.2	C 56
Sequence 226, App	5	US-09-978-360A-226	64.8	1283	16.2	C 57
Sequence 1325, A	7	US-10-995-561-1325	64.8	46089	16.2	C 58
Sequence 33, Appl	8	US-11-121-086-33	64.8	154548	16.2	C 59
Sequence 493604, A	25	US-11-121-849-493604	64.0	25	16	C 60
Sequence 62, Appl	77	US-10-310-914A-62	64.0	201	16	C 61
Sequence 48082, A	7	US-10-995-561-48082	64.0	201	16	C 62
Sequence 3068, Ap	8	US-11-043-752-3068	64.0	408	16	C 63
Sequence 1203, Ap	8	US-11-128-061-1203	64.0	580	16	C 64
Sequence 4845, Ap	8	US-11-128-061-4845	64.0	580	16	C 65
Sequence 1203, Ap	8	US-11-128-049-1203	64.0	580	16	C 66
Sequence 4845, Ap	8	US-11-128-049-4845	64.0	580	16	C 67
Sequence 4873, A	7	US-10-750-185-4873	64.0	754	16	C 68
Sequence 40873, A	7	US-10-750-623-40873	64.0	754	16	C 69
Sequence 32644, A	7	US-10-750-185-32644	64.0	787	16	C 70
Sequence 51825, A	7	US-10-750-623-51825	64.0	832	16	C 71
Sequence 51825, A	7	US-10-750-185-51825	64.0	832	16	C 72
Sequence 2345, Ap	7	US-10-750-623-2345	64.0	885	16	C 73
Sequence 34, Appl	6	US-10-981-334-34	64.0	999	16	C 74
Sequence 28120, A	7	US-10-750-185-28120	64.0	1249	16	C 75
Sequence 28120, A	7	US-10-750-623-28120	64.0	1249	16	C 76
Sequence 100, App	8	US-11-133-345-100	64.0	1557	16	C 77
Sequence 53359, A	7	US-10-750-185-53359	64.0	1665	16	C 78
Sequence 53359, A	7	US-10-750-623-53359	64.0	1665	16	C 79
Sequence 54888, A	7	US-10-750-185-54888	64.0	1669	16	C 80
Sequence 54888, A	7	US-10-750-623-54888	64.0	1669	16	C 81
Sequence 57434, A	7	US-10-750-185-57434	64.0	1726	16	C 82
Sequence 57434, A	7	US-10-750-623-57434	64.0	1726	16	C 83
Sequence 901, App	8	US-11-043-752-901	64.0	2001	16	C 84
Sequence 3469, Ap	7	US-10-793-626-3469	64.0	2489	16	C 85
Sequence 52920, A	7	US-10-750-185-52920	64.0	2670	16	C 86
Sequence 52920, A	7	US-10-750-623-52920	64.0	2670	16	C 87
Sequence 44806, A	7	US-10-750-185-44806	64.0	2824	16	C 88
Sequence 44806, A	7	US-10-750-623-44806	64.0	2824	16	C 89
Sequence 4170, Ap	7	US-10-793-626-4170	64.0	2902	16	C 90
Sequence 4170, Ap	7	US-10-793-626-4170	64.0	2902	16	C 91
Sequence 4120, Ap	7	US-10-793-626-4120	64.0	2929	16	C 92
Sequence 150, App	8	US-11-194-246-150	64.0	3000	16	C 93
Sequence 150, App	8	US-11-194-246-150	64.0	3000	16	C 94
Sequence 150, App	8	US-11-194-246-150	64.0	3000	16	C 95

c 96	16	64.0	3176	7	US-10-750-185-43645	Sequence 43645, A
c 97	16	64.0	3176	7	US-10-750-623-43645	Sequence 43645, A
c 98	16	64.0	3350	7	US-10-793-626-4018	Sequence 4018, Ap
c 99	16	64.0	3363	7	US-10-750-185-59149	Sequence 59149, A
c 100	16	64.0	3363	7	US-10-750-623-59149	Sequence 59149, A
c 101	16	64.0	3660	7	US-10-793-626-4116	Sequence 4116, Ap
c 102	16	64.0	3861	8	US-11-152-892-2	Sequence 2, Appl
c 103	16	64.0	3911	10	US-11-004-057-5	Sequence 5, Appl
c 104	16	64.0	4120	8	US-11-136-527-3046	Sequence 3046, Ap
c 105	16	64.0	59590	8	US-11-117-187-1847	Sequence 187, App
c 106	16	64.0	98309	8	US-11-124-368A-2921	Sequence 1821, Ap
c 107	16	64.0	151828	8	US-11-117-187-197	Sequence 197, App
c 108	16	64.0	155989	8	US-11-121-086-57	Sequence 57, Appl
c 109	16	64.0	170995	8	US-11-121-086-35	Sequence 35, Appl
c 110	16	64.0	178877	8	US-11-121-086-47	Sequence 17, Appl
c 111	16	64.0	180654	8	US-11-121-086-58	Sequence 58, Appl
c 112	16	64.0	184000	8	US-11-121-086-37	Sequence 37, Appl
c 113	16	64.0	645179	7	US-10-995-561-13293	Sequence 13293, A
c 114	16	64.0	1631140	8	US-11-091-018-1	Sequence 1, Appl
c 115	15.8	63.2	6040	7	US-10-240-708-69	Sequence 69, Appl
c 116	15.8	63.2	35100	8	US-11-127-832-26	Sequence 26, Appl
c 117	15.8	63.2	163289	8	US-11-121-086-20	Sequence 20, Appl
c 118	15.8	63.2	1080000	7	US-10-928-446A-1	Sequence 1, Appl
c 119	15.8	63.2	1080000	7	US-10-928-446A-181	Sequence 181, Appl
c 120	15.8	63.2	1080000	7	US-10-928-446A-183	Sequence 183, Appl
c 121	15.8	63.2	1080000	7	US-10-928-446A-185	Sequence 185, Appl
c 122	15.8	63.2	1080000	7	US-10-928-446A-187	Sequence 187, Appl
c 123	15.8	63.2	1080000	7	US-10-928-446A-189	Sequence 189, Appl
c 124	15.8	63.2	1080000	7	US-10-928-446A-191	Sequence 191, Appl
c 125	15.8	63.2	1080000	7	US-10-928-446A-193	Sequence 193, Appl
c 126	15.8	63.2	1080000	7	US-10-928-446A-195	Sequence 195, Appl
c 127	15.8	63.2	1080000	7	US-10-928-446A-197	Sequence 197, Appl
c 128	15.8	63.2	1080000	7	US-10-928-446A-199	Sequence 199, Appl
c 129	15.8	63.2	1080000	7	US-10-928-446A-201	Sequence 201, Appl
c 130	15.6	62.4	25	7	US-10-310-914A-148691	Sequence 148691, A
c 131	15.6	62.4	25	8	US-11-121-849-97549	Sequence 97549, A
c 132	15.6	62.4	201	7	US-10-995-561-57454	Sequence 57454, A
c 133	15.6	62.4	201	7	US-10-995-561-60293	Sequence 60293, A
c 134	15.6	62.4	201	7	US-10-995-561-76756	Sequence 76756, A
c 135	15.6	62.4	201	7	US-10-995-561-76969	Sequence 76969, A
c 136	15.6	62.4	817	7	US-10-750-185-50157	Sequence 50157, A
c 137	15.6	62.4	817	7	US-10-750-623-50157	Sequence 50157, A
c 138	15.6	62.4	918	8	US-11-074-176-289	Sequence 289, App
c 139	15.6	62.4	1073	7	US-10-750-185-31435	Sequence 31435, A
c 140	15.6	62.4	1073	7	US-10-750-623-31435	Sequence 31435, A
c 141	15.6	62.4	1107	7	US-10-750-185-59865	Sequence 59865, A
c 142	15.6	62.4	1107	7	US-10-750-623-59865	Sequence 59865, A
c 143	15.6	62.4	1382	7	US-10-750-185-39793	Sequence 39793, A
c 144	15.6	62.4	1382	7	US-10-750-623-39793	Sequence 39793, A
c 145	15.6	62.4	1887	7	US-10-793-626-2843	Sequence 2843, Ap
c 146	15.6	62.4	2112	7	US-10-821-234-560	

```

RESULT 1
US-11-136-527-6248
; Sequence 6248, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; CURRENT FILING DATE: 2005-03-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; PRIOR FILING DATE: 2005-05-26

```

```
Oy 3 AAAAAGTCATCCCAATGTTAT 24
Db 1441 AAAAAGTCATCCCTCAATGTTAT 1462

RESULT 4
US-10-793-626-4240
; Sequence 4240, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4240
; LENGTH: 3276
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-4240

Query Match 81.6%; Score 20.4; DB 7; Length 3276;
Best Local Similarity 95.5%; Pred. No. 3.2;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 3 AAAAAGTCATCCCAATGTTAT 24
Db 1193 AAAAAGTCATCCCTCAATGTTAT 1214

RESULT 5
US-10-857-780-7
; Sequence 7, Application US/10857780
; Publication No. US20050272043A1
; GENERAL INFORMATION:
; APPLICANT: ROTH, RICHARD B.
; APPLICANT: BRAUN, ANDREAS
; APPLICANT: KAMMERER, STEFAN M.
; APPLICANT: NELSON, MATTHEW ROBERTS
; APPLICANT: RENELAND, RIKARD HENRY
; APPLICANT: HOVAL-WRIGHTSON, CAROLYN R.
; TITLE OF INVENTION: METHODS FOR IDENTIFYING RISK OF BREAST CANCER AND TREATMENTS
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: SEQ-4069-CP
; CURRENT APPLICATION NUMBER: US/10/857,780
; CURRENT FILING DATE: 2004-05-28
; PRIOR APPLICATION NUMBER: 10/723,681
; PRIOR FILING DATE: 2003-11-25
; PRIOR APPLICATION NUMBER: 60/490,234
; PRIOR FILING DATE: 2003-07-24
; PRIOR APPLICATION NUMBER: 60/525,239
; PRIOR FILING DATE: 2003-11-25
; NUMBER OF SEQ ID NOS: 4962
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 7
; LENGTH: 95050
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (18543)..(18543)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (28737)..(28737)

OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (45763)..(45763)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (51178)..(51178)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (51180)..(51180)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (62241)..(62241)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (62541)..(62541)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (64042)..(64042)
OTHER INFORMATION: n is a, c, g, or t
US-10-857-780-7

Query Match 79.2%; Score 19.8; DB 7; Length 95050;
Best Local Similarity 91.3%; Pred. No. 12;
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 AAAAAGTCATCCCAATGTTA 23
Db 29898 AAACAAGTCACCCCAATGTTA 29920

RESULT 6
US-11-136-527-295557
; Sequence 295557, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 295557
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Probe
US-11-136-527-295557

Query Match 77.6%; Score 19.4; DB 8; Length 25;
Best Local Similarity 95.2%; Pred. No. 3.7;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 5 AACTGCATCCCAATGTTATG 25
Db 1 AACTGCATCCCAATGTTATG 21

RESULT 7
US-10-995-561-55567/c
; Sequence 55567, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
```


; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 48171
; LENGTH: 720
; TYPE: DNA
; ORGANISM: Bovine 19866880932834
US-10-750-623-48171

Query Match 70.4%; Score 17.6; DB 7; Length 720;
Best Local Similarity 83.3%; Pred. No. 50;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 AAAAACTGCATCCCAATGTTAT 24
Db 349 ATAAAACTGAATCCCAATGTTCT 372

RESULT 12
US-10-995-561-55566/c
; Sequence 55566, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 55566
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-55566

Query Match 68.8%; Score 17.2; DB 7; Length 201;
Best Local Similarity 79.2%; Pred. No. 60;
Matches 19; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 AAAAACTGCATCCCAATGTTAT 24
Db 102 AMCAAAATGCATCCCACTGTCAT 79

RESULT 13
US-11-124-368A-19759
; Sequence 19759, Application US/11124368A
; Publication No. US20050287559A1
; GENERAL INFORMATION:
; APPLICANT: Michele Cargill
; APPLICANT: James J. Devlin
; APPLICANT: May Luke
; TITLE OF INVENTION: Genetic Polymorphisms Associated with
; TITLE OF INVENTION: Vascular Diseases, Methods of Detection and Uses Thereof
; FILE REFERENCE: CL001524
; CURRENT APPLICATION NUMBER: US/11/124,368A
; CURRENT FILING DATE: 2005-05-09
; PRIOR APPLICATION NUMBER: US 60/568,845
; PRIOR FILING DATE: 2004-05-07
; PRIOR APPLICATION NUMBER: US 60/625,936
; PRIOR FILING DATE: 2004-11-09
; NUMBER OF SEQ ID NOS: 21112
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 19759
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-124-368A-19759

Query Match 68.8%; Score 17.2; DB 8; Length 201;

Best Local Similarity 86.4%; Pred. No. 60;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 2 AAAAACTGCATCCCAATGTTA 23
Db 101 AAAAACTTAATCCCAATGTGA 122

RESULT 14
US-11-124-368A-19877
; Sequence 19877, Application US/11124368A
; Publication No. US20050287559A1
; GENERAL INFORMATION:
; APPLICANT: Michele Cargill
; APPLICANT: James J. Devlin
; APPLICANT: May Luke
; TITLE OF INVENTION: Genetic Polymorphisms Associated with
; TITLE OF INVENTION: Vascular Diseases, Methods of Detection and Uses Thereof
; FILE REFERENCE: CL001524
; CURRENT APPLICATION NUMBER: US/11/124,368A
; CURRENT FILING DATE: 2005-05-09
; PRIOR APPLICATION NUMBER: US 60/568,845
; PRIOR FILING DATE: 2004-05-07
; PRIOR APPLICATION NUMBER: US 60/625,936
; PRIOR FILING DATE: 2004-11-09
; NUMBER OF SEQ ID NOS: 21112
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 19877
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-124-368A-19877

Query Match 68.8%; Score 17.2; DB 8; Length 201;
Best Local Similarity 86.4%; Pred. No. 60;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 AAAAACTGCATCCCAATGTTA 23
Db 132 AAAAACTTAATCCCAATGTGA 153

RESULT 15
US-10-750-185-1670
; Sequence 1670, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 1670
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Bovine MMBT19761
US-10-750-185-1670

Query Match 68.8%; Score 17.2; DB 7; Length 600;
Best Local Similarity 79.2%; Pred. No. 75;
Matches 19; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 2 AAAAACTGCATCCCAATGTTATG 25

Db 437 AMCACTGGACCCCAATTTATG 460
|: ||||| | ||||| |||||

RESULT 16
US-10-750-623-1670
; Sequence 1670, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1670
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Bovine MWBT19761
US-10-750-623-1670

Query Match 68.8%; Score 17.2; DB 7; Length 600;
Best Local Similarity 79.2%; Pred. No. 75;
Matches 19; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 AAAAACTGCATCCCAATGTTATG 25
|: ||||| | ||||| |||||
Db 437 AMCACTGGACCCCAATTTATG 460

RESULT 17
US-10-750-185-41308/c
; Sequence 41308, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 41308
; LENGTH: 855
; TYPE: DNA
; ORGANISM: Bovine 19866881259141
US-10-750-185-41308

Query Match 68.8%; Score 17.2; DB 7; Length 855;
Best Local Similarity 86.4%; Pred. No. 80;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAACTGCATCCCAATGTT 22
|: ||||| | ||||| |||||
Db 411 AAAAAATTCATCCCAATGTT 390

RESULT 18
US-10-750-623-41308/c
; Sequence 41308, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 41308
; LENGTH: 855
; TYPE: DNA
; ORGANISM: Bovine 19866881259141
US-10-750-623-41308

Query Match 68.8%; Score 17.2; DB 7; Length 855;
Best Local Similarity 86.4%; Pred. No. 80;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAACTGCATCCCAATGTT 22
|: ||||| | ||||| |||||
Db 411 AAAAAATTCATCCCAATGTT 390

RESULT 19
US-10-750-185-52076/c
; Sequence 52076, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 52076
; LENGTH: 1130
; TYPE: DNA
; ORGANISM: Bovine 19866880733723
US-10-750-185-52076

Query Match 68.8%; Score 17.2; DB 7; Length 1130;
Best Local Similarity 86.4%; Pred. No. 84;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAACTGCATCCCAATGTTA 23
|: ||||| | ||||| |||||
Db 85 AAAAAATTCATCCCAATTTTA 64

RESULT 20
US-10-750-623-52076/c

```
; Sequence 52076, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; PRIOR FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 52076
; LENGTH: 1130
; TYPE: DNA
; ORGANISM: Bovine 19866880733723
; US-10-750-623-52076

Query Match      68.8%; Score 17.2; DB 7; Length 1130;
Best Local Similarity 86.4%; Pred. No. 84;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 AAAAAATGCATCCCAATGTTA 23
Db 85 AAAAAATGCCTCCCAATTTA 64
||||| ||| ||||| |||

RESULT 21
US-11-124-368A-2924
; Sequence 2924, Application US/11124368A
; Publication No. US20050287559A1
; GENERAL INFORMATION:
; APPLICANT: Michele Cargill
; APPLICANT: James J. Devlin
; APPLICANT: May Luke
; TITLE OF INVENTION: Genetic Polymorphisms Associated with
; TITLE OF INVENTION: Vascular Diseases, Methods of Detection and Uses Thereof
; FILE REFERENCE: CL001524
; CURRENT APPLICATION NUMBER: US/11/124,368A
; CURRENT FILING DATE: 2005-05-09
; PRIOR APPLICATION NUMBER: US 60/568,845
; PRIOR FILING DATE: 2004-05-07
; PRIOR APPLICATION NUMBER: US 60/625,936
; PRIOR FILING DATE: 2004-11-09
; NUMBER OF SEQ ID NOS: 21112
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2924
; LENGTH: 79134
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-11-124-368A-2924

Query Match      68.8%; Score 17.2; DB 8; Length 79134;
Best Local Similarity 86.4%; Pred. No. 1.9e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 AAAAACTGAATCCCAATGTTA 23
Db 50290 AAAAACTGAATCCCAATGTTA 50311
||||| ||| ||||| |||

RESULT 22
US-11-121-086-31/c
; Sequence 31, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; APPLICANT: NIELSEN, KIRSTEN V.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138.6000-00000
; CURRENT APPLICATION NUMBER: US/11/121,086
; CURRENT FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 31
; LENGTH: 218821
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: modified_base
; LOCATION: (106949)..(106949)
; OTHER INFORMATION: a, c, g, t, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (110322)..(110324)
; OTHER INFORMATION: a, c, g, t, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (115133)..(115133)
; OTHER INFORMATION: a, c, g, t, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (131300)..(131300)
; OTHER INFORMATION: a, c, g, t, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (139059)..(139158)
; OTHER INFORMATION: a, c, g, t, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (157740)..(157740)
; OTHER INFORMATION: a, c, g, t, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (157777)..(157777)
; OTHER INFORMATION: a, c, g, t, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (157900)..(157900)
; OTHER INFORMATION: a, c, g, t, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (157919)..(157919)
; OTHER INFORMATION: a, c, g, t, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (157926)..(157926)
; OTHER INFORMATION: a, c, g, t, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (158094)..(158094)
; OTHER INFORMATION: a, c, g, t, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (158138)..(158138)
; OTHER INFORMATION: a, c, g, t, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (158193)..(158195)
; OTHER INFORMATION: a, c, g, t, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (158241)..(158242)
; OTHER INFORMATION: a, c, g, t, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (158259)..(158259)
; OTHER INFORMATION: a, c, g, t, unknown or other
```

```
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (158278)..(158278)
; OTHER INFORMATION: a, c, g, t, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (158295)..(158295)
; OTHER INFORMATION: a, c, g, t, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (158740)..(158839)
; OTHER INFORMATION: a, c, g, t, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (158929)..(158929)
; OTHER INFORMATION: a, c, g, t, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (163528)..(163530)
; OTHER INFORMATION: a, c, g, t, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (163550)..(163550)
; OTHER INFORMATION: a, c, g, t, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (163785)..(163785)
; OTHER INFORMATION: a, c, g, t, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (164000)..(164000)
; OTHER INFORMATION: a, c, g, t, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (164047)..(164047)
; OTHER INFORMATION: a, c, g, t, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (164084)..(164084)
; OTHER INFORMATION: a, c, g, t, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (167233)..(167233)
; OTHER INFORMATION: a, c, g, t, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (167236)..(167236)
; OTHER INFORMATION: a, c, g, t, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (167238)..(167238)
; OTHER INFORMATION: a, c, g, t, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (170247)..(170247)
; OTHER INFORMATION: a, c, g, t, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (170249)..(170250)
; OTHER INFORMATION: a, c, g, t, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (170252)..(170253)
; OTHER INFORMATION: a, c, g, t, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (170259)..(170259)
; OTHER INFORMATION: a, c, g, t, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (170262)..(170263)
; OTHER INFORMATION: a, c, g, t, unknown or other
; FEATURE:
```

```
; NAME/KEY: modified_base
; LOCATION: (170266)..(170266)
; OTHER INFORMATION: a, c, g, t, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (174470)..(174470)
; OTHER INFORMATION: a, c, g, t, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (174472)..(174472)
; OTHER INFORMATION: a, c, g, t, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (174474)..(174474)
; OTHER INFORMATION: a, c, g, t, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (179059)..(179060)
; OTHER INFORMATION: a, c, g, t, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (179064)..(179064)
; OTHER INFORMATION: a, c, g, t, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (197001)..(197001)
; OTHER INFORMATION: a, c, g, t, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (197004)..(197005)
; OTHER INFORMATION: a, c, g, t, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (197007)..(197007)
; OTHER INFORMATION: a, c, g, t, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (200349)..(200349)
; OTHER INFORMATION: a, c, g, t, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (200351)..(200351)
; OTHER INFORMATION: a, c, g, t, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (212425)..(212426)
; OTHER INFORMATION: a, c, g, t, unknown or other
; US-11-121-086-31
```

Query Match 68.8%; Score 17.2; DB 8; Length 218821;
Best Local Similarity 86.4%; Pred. No. 2.3e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AAAAACTGCATCCCAATGTT 22
|||||
Db 17443 AAAGACTGCATACCAATGTT 17422

RESULT 23
US-10-995-561-13327/c
; Sequence 13327, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13327

```
; LENGTH: 244196
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(244196)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-995-561-13327

Query Match      68.8%; Score 17.2; DB 7; Length 244196;
Best Local Similarity 79.2%; Pred. No. 2.3e+02; Indels 0; Gaps 0;
Matches 19; Conservative 1; Mismatches 4;

QY 1 AAAAACTGCATCCCAATGTTAT 24
    ||||| ||||| ||||| |||||
Db 28164 AMCAAAATGCATCCCACTGTAT 28141

RESULT 24
US-10-933-025-22/c
; Sequence 22, Application US/10933025
; Publication No. US20050265987A1
; GENERAL INFORMATION:
; APPLICANT: ROSEN, STEVEN
; APPLICANT: HEMMERICH, STEFAN
; APPLICANT: TOMITA, MEGUMI
; TITLE OF INVENTION: Sulfotransferases and methods of use
; FILE REFERENCE: UCAL-230CON
; CURRENT APPLICATION NUMBER: US/10/933,025
; CURRENT FILING DATE: 2004-09-01
; PRIOR APPLICATION NUMBER: 10/025,966
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/258,577
; PRIOR FILING DATE: 2000-12-27
; PRIOR APPLICATION NUMBER: 60/267,831
; PRIOR FILING DATE: 2001-09-02
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 268685
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(268685)
; OTHER INFORMATION: n = A,T,C or G
US-10-933-025-22

Query Match      68.8%; Score 17.2; DB 7; Length 268685;
Best Local Similarity 86.4%; Pred. No. 2.3e+02; Indels 0; Gaps 0;
Matches 19; Conservative 0; Mismatches 3;

QY 3 AAAAACTGCATCCCAATGTTAT 24
    ||||| ||||| ||||| |||||
Db 265891 AATACTGCATCCGCAATGTAT 265870

RESULT 25
US-10-750-185-827
; Sequence 827, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
```

```
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 827
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Bovine MMBT09456
US-10-750-185-827

Query Match      68.0%; Score 17; DB 7; Length 600;
Best Local Similarity 80.0%; Pred. No. 93;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAACTGCATCCCAATGTTATG 25
    ||||| ||||| ||||| |||||
Db 176 AAAAAATCATCCCTTAAGTTATG 200

RESULT 26
US-10-750-623-827
; Sequence 827, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 827
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Bovine MMBT09456
US-10-750-623-827

Query Match      68.0%; Score 17; DB 7; Length 600;
Best Local Similarity 80.0%; Pred. No. 93;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAACTGCATCCCAATGTTATG 25
    ||||| ||||| ||||| |||||
Db 176 AAAAAATCATCCCTTAAGTTATG 200

RESULT 27
US-11-121-086-73
; Sequence 73, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; APPLICANT: NIELSEN, KIRSTEN V.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138.6000-00000
; CURRENT APPLICATION NUMBER: US/11/121,086
; CURRENT FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 73
; LENGTH: 152335
; TYPE: DNA
```

```
; ORGANISM: Homo sapiens
US-11-121-086-73

Query Match      68.0%; Score 17; DB 8; Length 152335;
Best Local Similarity 80.0%; Pred. No. 2.6e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAACTGCATCCCAATGTTATG 25
   ||||| ||||| ||||| ||||| |||||
Db 150782 AACACTTGAAATCCCAATGTTATG 150806

RESULT 28
US-10-995-561-13233
; Sequence 13233, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13233
; LENGTH: 199130
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; LOCATION: (1)...(199130)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-995-561-13233

Query Match      68.0%; Score 17; DB 7; Length 199130;
Best Local Similarity 80.0%; Pred. No. 2.8e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAACTGCATCCCAATGTTATG 25
   ||||| ||||| ||||| ||||| |||||
Db 56718 AACATCTGCATCCCAATGTTATG 56742

RESULT 29
US-11-102-978-3
; Sequence 3, Application US/11102978
; Publication No. US20050250142A1
; GENERAL INFORMATION:
; APPLICANT: University of Utah Technology Transfer Office
; APPLICANT: University of Utah Research Foundation
; TITLE OF INVENTION: Diagnosis and Treatment of Herpes Simplex Virus Disease
; FILE REFERENCE: 0274-5537-1US
; CURRENT APPLICATION NUMBER: US/11/102,978
; CURRENT FILING DATE: 2005-04-11
; PRIOR APPLICATION NUMBER: PCT/US2003/033152
; PRIOR FILING DATE: 2003-10-18
; PRIOR APPLICATION NUMBER: 60/419,576
; PRIOR FILING DATE: 2002-10-18
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3
; LENGTH: 340000
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: exon
; LOCATION: (56948)..(57115)
; OTHER INFORMATION: C21orf34 exon
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (80006)..(81089)
```

```
; OTHER INFORMATION: Gene VDACP; voltage-dependent anion channel isoform 2 pseudogene
; FEATURE:
; NAME/KEY: exon
; LOCATION: (167308)..(167438)
; OTHER INFORMATION: C21orf34 exon
; FEATURE:
; NAME/KEY: exon
; LOCATION: (216732)..(216833)
; OTHER INFORMATION: C21orf34 exon
US-11-102-978-3

Query Match      68.0%; Score 17; DB 8; Length 340000;
Best Local Similarity 80.0%; Pred. No. 3e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAACTGCATCCCAATGTTATG 25
   ||||| ||||| ||||| ||||| |||||
Db 182887 AAAAAATGCTTACCAATGGTGTG 182911

RESULT 30
US-11-121-086-64/c
; Sequence 64, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138.6000-00000
; CURRENT APPLICATION NUMBER: US/11/121,086
; CURRENT FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 64
; LENGTH: 142605
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-64

Query Match      67.2%; Score 16.8; DB 8; Length 142605;
Best Local Similarity 90.0%; Pred. No. 3.2e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 AAATGCAATCCCAATGTTA 23
   ||||| ||||| ||||| ||||| |||||
Db 26350 AAATCCATCCCAATGTGA 26331

RESULT 31
US-11-121-086-63/c
; Sequence 63, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138.6000-00000
; CURRENT APPLICATION NUMBER: US/11/121,086
; CURRENT FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 63
; LENGTH: 169725
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-63

Query Match      67.2%; Score 16.8; DB 8; Length 169725;
Best Local Similarity 90.0%; Pred. No. 3.3e+02;
```

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 AACTGCATCCCAATGTTA 23
||||| ||||| ||||| ||||| |||||
Db 168192 AACTCATCCCAATGTGA 168173

RESULT 32

US-10-995-561-13215
; Sequence 13215, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13215
; LENGTH: 241805
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(241805)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-995-561-13215

Query Match 67.2%; Score 16.8; DB 7; Length 241805;
Best Local Similarity 90.0%; Pred. No. 3.5e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 AACTGCATCCCAATGTTAT 24
||||| ||||| ||||| ||||| |||||
Db 78480 AACTGCAACCCCAAGTTAT 78499

RESULT 33

US-10-750-185-19474/c
; Sequence 19474, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19474
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Bovine MMBT04310
US-10-750-185-19474

Query Match 66.4%; Score 16.6; DB 7; Length 600;
Best Local Similarity 82.6%; Pred. No. 1.4e+02;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AAAAAGCGCATCCCAATGTTAT 24
||||| ||||| ||||| ||||| |||||
Db 568 AAAAGCGCATCCCAAGATAT 546

RESULT 34

US-10-750-623-19474/c
; Sequence 19474, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19474
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Bovine MMBT04310
US-10-750-623-19474

Query Match 66.4%; Score 16.6; DB 7; Length 600;
Best Local Similarity 82.6%; Pred. No. 1.4e+02;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AAAAAGCGCATCCCAATGTTAT 24
||||| ||||| ||||| ||||| |||||
Db 568 AAAAGCGCATCCCAAGATAT 546

RESULT 35

US-10-750-185-24537/c
; Sequence 24537, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 24537
; LENGTH: 1463
; TYPE: DNA
; ORGANISM: Bovine 19866880525178
US-10-750-185-24537

Query Match 66.4%; Score 16.6; DB 7; Length 1463;
Best Local Similarity 82.6%; Pred. No. 1.7e+02;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AAAAAGCGCATCCCAATGTTAT 24
||||| ||||| ||||| ||||| |||||
Db 1070 AAAAGCGCATCCCAAGATAT 1048

RESULT 36

US-10-750-623-24537/c
; Sequence 24537, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 24537
; LENGTH: 1463
; TYPE: DNA
; ORGANISM: Bovine 19866880525178
US-10-750-623-24537
Query Match 66.4%; Score 16.6; DB 7; Length 1463;
Best Local Similarity 82.6%; Pred. No. 1.7e+02;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 2 AAAAAGCGCATCCCCAATGTTAT 24
||||| ||||| ||||| ||||| |||||
Db 1070 AAAAGCGCATCCCCAAGATAT 1048
||||| ||||| ||||| ||||| |||||
RESULT 37
US-11-033-764-94/c
; Sequence 94, Application US/11033764
; Publication No. US20050244817A1
; GENERAL INFORMATION:
; APPLICANT: Rubin, Donald H.
; APPLICANT: Organ, Edward L.
; APPLICANT: DuBois, Raymond N.
; TITLE OF INVENTION: Mammalian Genes Involved in Viral
; TITLE OF INVENTION: Infection and Tumor Suppression
; FILE REFERENCE: 01123.0004
; CURRENT APPLICATION NUMBER: US/11/033,764
; CURRENT FILING DATE: 2005-01-12
; PRIOR APPLICATION NUMBER: US/10/877,807
; PRIOR FILING DATE: 2004-06-25
; PRIOR APPLICATION NUMBER: US/09/509,712
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: PCT/US98/21276
; PRIOR FILING DATE: 1998-10-08
; PRIOR APPLICATION NUMBER: 60/062,021
; PRIOR FILING DATE: 1997-10-10
; NUMBER OF SEQ ID NOS: 127
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 94
; LENGTH: 2161
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1- 2161
; OTHER INFORMATION: n = g, a, c or t(u)
US-11-033-764-94
Query Match 66.4%; Score 16.6; DB 10; Length 2161;
Best Local Similarity 82.6%; Pred. No. 1.8e+02;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 3 AAAAAGCGCATCCCCAATGTTATG 25
||||| ||||| ||||| ||||| |||||

Db 1647 AAAAAGCGCATCCCCCTACTGTTCTG 1625
RESULT 38
US-10-750-185-29455/c
; Sequence 29455, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 29455
; LENGTH: 2524
; TYPE: DNA
; ORGANISM: Bovine 19866880652511
US-10-750-185-29455
Query Match 66.4%; Score 16.6; DB 7; Length 2524;
Best Local Similarity 82.6%; Pred. No. 1.9e+02;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 2 AAAAAGCGCATCCCCAATGTTAT 24
||||| ||||| ||||| ||||| |||||
Db 2027 AAAAAGCGCATGTAATGTTAT 2005
||||| ||||| ||||| ||||| |||||
RESULT 39
US-10-750-623-29455/c
; Sequence 29455, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 29455
; LENGTH: 2524
; TYPE: DNA
; ORGANISM: Bovine 19866880652511
US-10-750-623-29455
Query Match 66.4%; Score 16.6; DB 7; Length 2524;
Best Local Similarity 82.6%; Pred. No. 1.9e+02;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 2 AAAAAGCGCATCCCCAATGTTAT 24
||||| ||||| ||||| ||||| |||||
Db 2027 AAAAAGCGCATGTAATGTTAT 2005
||||| ||||| ||||| ||||| |||||


```
RESULT 42
US-11-112-908-63
; Sequence 63, Application US/11112908
; Publication No. US20050260659A1
; GENERAL INFORMATION:
; APPLICANT: Harris, Cole
; APPLICANT: Davis, Lisa M.
; TITLE OF INVENTION: Breast Cancer Biomarkers
; FILE REFERENCE: 04-164-US
; CURRENT APPLICATION NUMBER: US/11/112,908
; CURRENT FILING DATE: 2005-04-22
; PRIOR APPLICATION NUMBER: US 60/564,758
; PRIOR FILING DATE: 2004-04-23
; PRIOR APPLICATION NUMBER: US 60/575,978
; PRIOR FILING DATE: 2004-06-01
; PRIOR APPLICATION NUMBER: US 60/631,702
; PRIOR FILING DATE: 2004-11-30
; PRIOR APPLICATION NUMBER: US 60/633,826
; PRIOR FILING DATE: 2004-12-07
; NUMBER OF SEQ ID NOS: 511
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 63
; TYPE: DNA
; LENGTH: 149111
; ORGANISM: Homo sapiens
; LOCATION: (162)...(5363)
US-11-112-908-63

Query Match 66.4%; Score 16.6; DB 8; Length 149111;
Best Local Similarity 82.6%; Pred. No. 4e+02;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAACTGCATCCCAATGTTA 23
Db 19139 AAAAACTGCATCCCAATGTTA 19161

RESULT 43
US-11-112-908-61
; Sequence 61, Application US/11112908
; Publication No. US20050260659A1
; GENERAL INFORMATION:
; APPLICANT: Harris, Cole
; APPLICANT: Davis, Lisa M.
; TITLE OF INVENTION: Breast Cancer Biomarkers
; FILE REFERENCE: 04-164-US
; CURRENT APPLICATION NUMBER: US/11/112,908
; CURRENT FILING DATE: 2005-04-22
; PRIOR APPLICATION NUMBER: US 60/564,758
; PRIOR FILING DATE: 2004-04-23
; PRIOR APPLICATION NUMBER: US 60/575,978
; PRIOR FILING DATE: 2004-06-01
; PRIOR APPLICATION NUMBER: US 60/631,702
; PRIOR FILING DATE: 2004-11-30
; PRIOR APPLICATION NUMBER: US 60/633,826
; PRIOR FILING DATE: 2004-12-07
; NUMBER OF SEQ ID NOS: 511
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 61
; TYPE: DNA
; LENGTH: 159497
; ORGANISM: Homo sapiens
; LOCATION: (162)...(5363)
US-11-112-908-61

Query Match 66.4%; Score 16.6; DB 8; Length 159497;
Best Local Similarity 82.6%; Pred. No. 4e+02;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAACTGCATCCCAATGTTA 23
Db 70881 AAAAACTGCAGGCAAAATGTTA 70903
```

```
RESULT 40
US-11-192-967-5/c
; Sequence 5, Application US/11192967
; Publication No. US20050262587A1
; GENERAL INFORMATION:
; APPLICANT: Flannagan, Ronald D.
; APPLICANT: Mathis, John P.
; APPLICANT: Meyer, Terry E.
; TITLE OF INVENTION: Novel Bt Toxin Receptors From
; FILE REFERENCE: 35718/204664
; CURRENT APPLICATION NUMBER: US/11/192,967
; CURRENT FILING DATE: 2005-07-29
; PRIOR APPLICATION NUMBER: US/09/715,909
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/166,285
; PRIOR FILING DATE: 1999-11-18
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 5592
; TYPE: DNA
; ORGANISM: Spodoptera frugiperda
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (162)...(5363)
US-11-192-967-5

Query Match 66.4%; Score 16.6; DB 8; Length 5592;
Best Local Similarity 82.6%; Pred. No. 2.2e+02;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AAAAACTGCATCCCAATGTTAT 24
Db 1700 AATAATTGCTTCCCAATGTGAT 1678

RESULT 41
US-11-193-715-5/c
; Sequence 5, Application US/11193715
; Publication No. US20050260675A1
; GENERAL INFORMATION:
; APPLICANT: Flannagan, Ronald D.
; APPLICANT: Mathis, John P.
; APPLICANT: Meyer, Terry E.
; TITLE OF INVENTION: Novel Bt Toxin Receptors From
; FILE REFERENCE: 35718/204664
; CURRENT APPLICATION NUMBER: US/11/193,715
; CURRENT FILING DATE: 2005-07-29
; PRIOR APPLICATION NUMBER: US/09/715,909
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/166,285
; PRIOR FILING DATE: 1999-11-18
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 5592
; TYPE: DNA
; ORGANISM: Spodoptera frugiperda
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (162)...(5363)
US-11-193-715-5

Query Match 66.4%; Score 16.6; DB 8; Length 5592;
Best Local Similarity 82.6%; Pred. No. 2.2e+02;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AAAAACTGCATCCCAATGTTAT 24
Db 1700 AATAATTGCTTCCCAATGTGAT 1678
```

RESULT 44
US-11-121-086-44
; Sequence 44, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; APPLICANT: NIELSEN, KIRSTEN V.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138.6000-00000
; CURRENT APPLICATION NUMBER: US/11/121,086
; CURRENT FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 44
; LENGTH: 167116
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-44

Query Match 66.4%; Score 16.6; DB 8; Length 167116;
Best Local Similarity 82.6%; Pred. No. 4.1e+02;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 AAAAACTGCATCCCAATGTTA 23
DB 1917 AGAAATCTGTATCCCAATGTTA 1939

RESULT 45
US-11-121-086-43
; Sequence 43, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; APPLICANT: NIELSEN, KIRSTEN V.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138.6000-00000
; CURRENT APPLICATION NUMBER: US/11/121,086
; CURRENT FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 43
; LENGTH: 175416
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-43

Query Match 66.4%; Score 16.6; DB 8; Length 175416;
Best Local Similarity 82.6%; Pred. No. 4.1e+02;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 AAAAACTGCATCCCAATGTTA 23
DB 168998 AGAAATCTGTATCCCAATGTTA 169020

RESULT 46
US-11-112-908-20/c
; Sequence 20, Application US/11112908
; Publication No. US20050260659A1
; GENERAL INFORMATION:
; APPLICANT: Harris, Cole
; APPLICANT: Davis, Lisa M.
; TITLE OF INVENTION: Breast Cancer Biomarkers
; FILE REFERENCE: 04-164-US
; CURRENT APPLICATION NUMBER: US/11/112,908
; CURRENT FILING DATE: 2005-04-22
; PRIOR APPLICATION NUMBER: US 60/564,758
; PRIOR FILING DATE: 2004-04-23

; PRIOR APPLICATION NUMBER: US 60/575,978
; PRIOR FILING DATE: 2004-06-01
; PRIOR APPLICATION NUMBER: US 60/631,702
; PRIOR FILING DATE: 2004-11-30
; PRIOR APPLICATION NUMBER: US 60/633,826
; PRIOR FILING DATE: 2004-12-07
; NUMBER OF SEQ ID NOS: 511
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 20
; LENGTH: 191331
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-112-908-20

Query Match 66.4%; Score 16.6; DB 8; Length 191331;
Best Local Similarity 82.6%; Pred. No. 4.2e+02;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 2 AAAAACTGCATCCCAATGTTAT 24
DB 4946 AAAAAATTTCTCCCAATATAT 4924

RESULT 47
US-10-750-185-56163/c
; Sequence 56163, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 56163
; LENGTH: 1299
; TYPE: DNA
; ORGANISM: Bovine 19866880487873
US-10-750-185-56163

Query Match 65.6%; Score 16.4; DB 7; Length 1299;
Best Local Similarity 94.4%; Pred. No. 2.1e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 AAAAAACTGCATCCCAAA 18
DB 164 AAAAACTGAATCCCAAA 147

RESULT 48
US-10-750-623-56163/c
; Sequence 56163, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 AAAGTGCATCCCCAATGT 21
||||| |||||||
Db 134276 AAAGTGCATCCCCAATGT 134293

Search completed: February 3, 2006, 16:20:17
Job time : 353.111 secs

;; CURRENT FILING DATE: 2003-12-31
;; PRIOR APPLICATION NUMBER: US 60/437,482
;; PRIOR FILING DATE: 2002-12-31
;; NUMBER OF SEQ ID NOS: 64922
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 56163
;; LENGTH: 1299
;; TYPE: DNA
;; ORGANISM: Bovine 19866880487873
US-10-750-623-56163

Query Match 65.6%; Score 16.4; DB 7; Length 1299;
Best Local Similarity 94.4%; Pred. No. 2.1e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AAAGTGCATCCCCAA 18
||||| |||||||
Db 164 AAAGTGCATCCCCAA 147

RESULT 49
US-11-136-527-2599/c
; Sequence 2599, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2599
; LENGTH: 2217
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-2599

Query Match 65.6%; Score 16.4; DB 8; Length 2217;
Best Local Similarity 94.4%; Pred. No. 2.3e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 7 CTCATCCCCCAATGTTAT 24
||||| |||||||
Db 226 CTCATCCCCCAATGTTAT 209

RESULT 50
US-11-121-086-39
; Sequence 39, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; APPLICANT: NIELSEN, KIRSTEN V.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138.6000-00000
; CURRENT APPLICATION NUMBER: US/11/121,086
; CURRENT FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 39
; LENGTH: 207835
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-39

Query Match 65.6%; Score 16.4; DB 8; Length 207835;
Best Local Similarity 94.4%; Pred. No. 5.1e+02;

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 3, 2006, 10:53:23 ; Search time 863 Seconds
(without alignments)
1646.682 Million cell updates/sec

Title: US-10-719-900-30
Perfect score: 25
Sequence: 1 aaaaactggaatctcagctgaga 25

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database :

GenEmbl:*

- 1: gb_ba:*
- 2: gb_in:*
- 3: gb_env:*
- 4: gb_ov:*
- 5: gb_ov:*
- 6: gb_pat:*
- 7: gb_ph:*
- 8: gb_pr:*
- 9: gb_ro:*
- 10: gb_sts:*
- 11: gb_sy:*
- 12: gb_un:*
- 13: gb_vi:*
- 14: gb_hcg:*
- 15: gb_pl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	25	100.0	5223	9	AK173165 Mus muscu
C 2	25	100.0	195988	9	AC092481 Mus muscu
C 3	25	100.0	213144	9	AC092480 Mus muscu
4	25	100.0	214356	14	AC154650 Mus muscu
5	25	100.0	243637	14	AC097150 Rattus no
6	25	100.0	248916	14	AC137217 Rattus no
C 7	20.2	80.8	25837	14	AC087198 Homo sapi
C 8	20.2	80.8	30005	8	AL596124 Human DNA
C 9	20.2	80.8	180569	8	AL354949 Human DNA
C 10	20.2	80.8	193379	14	AC156206 Bos tauru
C 11	20.2	80.8	194019	14	AC053464 Homo sapi
C 12	20.2	80.8	208202	14	AC013446 Homo sapi
C 13	20.2	80.8	231342	14	AC152756 Bos tauru
14	20	80.0	300	6	BD212075 Novel hum
15	20	80.0	758	6	BD215003 Novel hum
16	20	80.0	2272	8	AB168750 Macaca fa
17	20	80.0	4810	8	AB037818 Homo sapi
18	20	80.0	5174	6	AX405807 Sequence

C 92	18.6	74.4	850	15	DCU52848	U52848 Daucus caro
C 93	18.6	74.4	1000	10	BV545071	BV545071 sxe63D03.
C 94	18.6	74.4	1047	4	AB198070	AB198070 Bos tauru
C 95	18.6	74.4	1364	8	BSM807436	BSM807436 Homo sapi
C 96	18.6	74.4	1388	8	AK126147	AK126147 Homo sapi
C 97	18.6	74.4	2066	6	AX713787	AX713787 Sequence
C 98	18.6	74.4	2066	8	AK055576	AK055576 Homo sapi
C 99	18.6	74.4	4527	15	PANMTC01	M28703 Podospora a
C 100	18.6	74.4	29227	2	AF000192	AF000192 Caenorhab
C 101	18.6	74.4	30223	1	AE001576	AE001576 Borrelia
C 102	18.6	74.4	30316	8	AC124219	AC124219 Pan trogl
C 103	18.6	74.4	40179	8	AC004036	AC004036 Homo sapi
C 104	18.6	74.4	44788	8	AL133267	AL133267 Human DNA
C 105	18.6	74.4	45878	15	AB016883	AB016883 Arabidops
C 106	18.6	74.4	52371	1	AE001584	AE001584 Borrelia
C 107	18.6	74.4	54398	14	AC068224	AC068224 Homo sapi
C 108	18.6	74.4	63771	8	AL590425	AL590425 Human DNA
C 109	18.6	74.4	68581	8	AP000563	AP000563 Homo sapi
C 110	18.6	74.4	80682	9	AL953839	AL953839 Mouse DNA
C 111	18.6	74.4	82789	14	AC015629	AC015629 Homo sapi
C 112	18.6	74.4	95477	8	AC007076	AC007076 Homo sapi
C 113	18.6	74.4	97777	14	AC165650	AC165650 Bos tauru
C 114	18.6	74.4	99314	14	AC084061	AC084061 Mus muscu
C 115	18.6	74.4	100000	8	AP000211	AP000211 Homo sapi
C 116	18.6	74.4	100314	15	MTPACG	XS5026 Podospora a
C 117	18.6	74.4	100348	8	AC091860	AC091860 Homo sapi
C 118	18.6	74.4	100960	14	AC166764	AC166764 Bos tauru
C 119	18.6	74.4	102649	14	AC134715	AC134715 Rattus no
C 120	18.6	74.4	110000	1	AE006470.14	Continuation (15 o
C 121	18.6	74.4	110000	14	BX324168.0	BX324168 Homo sapi
C 122	18.6	74.4	110000	14	TANN2_00	CR940348 Theileria
C 123	18.6	74.4	122001	8	AC007215	AC007215 Homo sapi
C 124	18.6	74.4	125066	8	AC074286	AC074286 Homo sapi
C 125	18.6	74.4	126852	8	AC010478	AC010478 Homo sapi
C 126	18.6	74.4	127675	8	AC146221	AC146221 Pan trogl
C 127	18.6	74.4	130158	14	AC166075	AC166075 Mus muscu
C 128	18.6	74.4	132890	14	AC163197	AC163197 Bos tauru
C 129	18.6	74.4	135146	14	AC161719	AC161719 Loxodonta
C 130	18.6	74.4	137246	8	HS49J10	Z84572 Human DNA s
C 131	18.6	74.4	139719	14	AC141439	AC141439 Homo sapi
C 132	18.6	74.4	139744	5	BX957315	BX957315 Zebrafish
C 133	18.6	74.4	142948	8	AC147025	AC147025 Pan trogl
C 134	18.6	74.4	146259	8	AL137013	AL137013 Human DNA
C 135	18.6	74.4	149965	8	AC007718	AC007718 Homo sapi
C 136	18.6	74.4	152311	14	AC011571	AC011571 Homo sapi
C 137	18.6	74.4	160606	8	AC093517	AC093517 Homo sapi
C 138	18.6	74.4	161781	14	CR376800	CR376800 Danio rer
C 139	18.6	74.4	162377	8	AL683870	AL683870 Human DNA
C 140	18.6	74.4	163046	9	AC129585	AC129585 Mus muscu
C 141	18.6	74.4	164294	14	AC148910	AC148910 Otolemur
C 142	18.6	74.4	166093	9	AC119971	AC119971 Mus muscu
C 143	18.6	74.4	168487	8	AL138820	AL138820 Human DNA
C 144	18.6	74.4	168799	8	AC009531	AC009531 Homo sapi
C 145	18.6	74.4	169786	8	BS000112	BS000112 Pan trogl
C 146	18.6	74.4	170654	8	AC073901	AC073901 Homo sapi
C 147	18.6	74.4	173021	8	AC146376	AC146376 Homo sapi
C 148	18.6	74.4	176277	14	AC025941	AC025941 Pan trogl
C 149	18.6	74.4	179906	14	AC119339	AC119339 Rattus no
C 150	18.6	74.4	181446	14	CR954266	CR954266 Danio rer

ALIGNMENTS

RESULT 1	AK173165	AK173165	5223 bp	mRNA	linear	ROD 28-JUL-2004
LOCUS	Mus musculus mRNA for mKIAA1397 protein.					
DEFINITION	AK173165					
ACCESSION	AK173165					
VERSION	AK173165.1	GI:50510914				
KEYWORDS	FLI_CDNA.					
SOURCE	Mus musculus (house mouse)					
ORGANISM	Mus musculus					

REFERENCE	1	Okazaki,N., Kikuno,R.F., Ohara,R., Inamoto,S., Koseki,H., Hiraoka,S., Saga,Y., Seino,S., Nishimura,M., Kaisho,T., Hoshino,K., Kitamura,H., Nagase,T., Ohara,O. and Koga,H.	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE		Prediction of the Coding Sequences of Mouse Homologues of KIAA	
JOURNAL		Gene: IV. The Complete Nucleotide Sequences of 500 Mouse KIAA-Homologous cDNAs Identified by Screening of Terminal Sequences of cDNA Clones Randomly Sampled from Size-Fractionated Libraries	
REFERENCE	2	(bases 1 to 5223)	
AUTHORS	Okazaki,N., Kikuno,R.F., Nagase,T., Ohara,O. and Koga,H.		
TITLE	Direct Submission		
JOURNAL	Submitted (19-MAY-2004) Hisashi Koga, Kazusa DNA Research Institute, Laboratory for Genome Informatics; 2-6-7 Kazusa-kamatari, Kisarazu, Chiba 292-0818, Japan (E-mail:mousse@kazusa.or.jp, Tel:81-438-52-3919, Fax:81-438-52-3918)		
COMMENT	The CREATE program supported by Japan science and technology corporation; cDNA full insert sequencing; Kazusa DNA Research Institute; cDNA library construction, clone selection and 5'- & 3'-end one pass sequencing.		
FEATURES	Location/Qualifiers		
source	1. 5223		
	/organism="Mus musculus"		
	/mol_type="mRNA"		
	/db_xref="taxon:10090"		
	/clone="mbgl12832"		
	/tissue_type="brain"		
	/dev_stage="adult"		
	note="vector:modified pBC SK+"		
gene	1. 5223		
	/gene="mKIAA1397"		
CDS	<623..2833		
	/gene="mKIAA1397"		
	/note="CDS is predicted by in silico analysis. Start codon is not identified."		
	/codon_start=1		
	/evidence=not experimental		
	/product="mKIAA1397 protein"		
	/protein_id="BAD32443.1"		
	/db_xref="GI:50510915"		
	/translation="LDSRSVYTKKKNKPTPNHKNKDKIKELGLKFLSTFKVFKVLEK SMYAAVEHPVLCDSDNLCUSKGRVPEKSEKPCVRRRIYEGNLAIGRQGVGV TFTSHCRDRSTPQFINFNLRHNSVYLRWNEPYOKLATCDAGGIFVMIQYGR WSVELNDRGAQVSDFTWSDGTQALISYRDGFVLVSGVQQRHWSINLESQITCG IWTDDQVLFGTADGOVIMDCRGLAHVLLHESDGLSWMYNYPIFLVEDSSESD TSDSDYSPDGPAAYPPIPVONTKPLLTVSFTSGDISLMNNYDDLSPVIRSGLKEV AQWCTQGLLAVAGNEQAQSELNPGELLKSAWYKFTYVNRGEHIFTLDTLVQRPILS ICWHRDRSRLWASGPALYVVRVHRVSSQLLCCQAIASTRDKDVNKTLPRLC SYLTAFTPTTKPIPPDPNNRDRFVSPSAGNERLHCTMKRTDDPFGVGGPYTLLE YLGLVPTLTKGRISKLKPEPVIDPRTDSKSDIYGNLSLSTVDSNCSDSDIEL SDDWAAKSPKISRSKSPKLPRISEARKSPKLPRAAQEI SRSPRLPMKPSMGSPS LTRREFPEDITOHNYLAQVTSNIGTWGTFKIVGLAAPIPTNLGACKDSLIFLTVPLEI NKSSHLNQLQFAFENEFLRQRTPTVCLLDDSSHLSLEKQLCRSLALFLELSCVFL"		
ORIGIN			
Query Match	100.0%;	Score 25;	DB 9; Length 5223;
Best Local Similarity	100.0%;	Pred. No. 0.47;	
Matches	25; Conservative	0; Mismatches	0; Indels 0; Gaps 0;
OY	1	AAAAAATCGGAATCTCAGGCTGAGA	25
Db	381	AAAAAATCGGAATCTCAGGCTGAGA	405
RESULT 2	AC092481/c		
LOCUS	AC092481/c		
DEFINITION	Mus musculus strain C57BL/6J chromosome 17 clone rp23-456p8, complete sequence.		
ACCESSION	AC092481		

```

VERSION AC092481.28 GI:25013385
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
TITLE Sciurognathi; Muridae; Murinae; Mus.
JOURNAL 1 (bases 1 to 195988)
REFERENCE Li, S., Tian, R. and Roe, B.A.
AUTHORS Mus musculus BAC Clone rp23-456p8
TITLE Unpublished
JOURNAL 2 (bases 1 to 195988)
REFERENCE Li, S., Tian, R. and Roe, B.A.
AUTHORS Direct Submission
TITLE Submitted (10-JUL-2001) Department of Chemistry And Biochemistry,
JOURNAL The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
REFERENCE 3 (bases 1 to 195988)
AUTHORS Li, S., Tian, R. and Roe, B.A.
TITLE Direct Submission
JOURNAL Submitted (26-APR-2002) Department of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
REFERENCE 4 (bases 1 to 195988)
AUTHORS Li, S., Tian, R. and Roe, B.A.
TITLE Direct Submission
JOURNAL Submitted (03-MAY-2002) Department of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
REFERENCE 5 (bases 1 to 195988)
AUTHORS Li, S., Tian, R. and Roe, B.A.
TITLE Direct Submission
JOURNAL Submitted (04-MAY-2002) Department of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
REFERENCE 6 (bases 1 to 195988)
AUTHORS Li, S., Tian, R. and Roe, B.A.
TITLE Direct Submission
JOURNAL Submitted (16-JUL-2002) Department of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
REFERENCE 7 (bases 1 to 195988)
AUTHORS Li, S., Tian, R. and Roe, B.A.
TITLE Direct Submission
JOURNAL Submitted (15-NOV-2002) Department of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
COMMENT On Nov 15, 2002 this sequence version replaced gi:21844618.
----- Genome Center
Center: Department Of Chemistry And Biochemistry
The University Of Oklahoma
Center code:UOKNOR

FEATURES             Location/Qualifiers
     source            1..195988
                        /organism="Mus musculus"
                        /mol_type="genomic DNA"
                        /strain="C57BL/6J"
                        /db_xref="taxon:10090"
                        /chromosome="17"
                        /clone="rp23-456p8"
                        /clone_lib="RPCI - 23 Female (C57BL/6J) Mouse BAC Library"

ORIGIN
Query Match      100.0%; Score 25; DB 9; Length 195988;
Best Local Similarity 100.0%; Pred. No. 0.38;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAAAACTGGAATCTCAGGCTGAGA 25
    |||||
Db 121595 AAAAACTGGAATCTCAGGCTGAGA 121571

RESULT 3
AC092480/c          213144 bp      DNA      linear      ROD 15-NOV-2002
LOCUS              Mus musculus strain C57BL/6J chromosome 17 clone rp23-387f2,
DEFINITION          complete sequence.
ACCESSION           AC092480
VERSION             AC092480.27 GI:25013374
KEYWORDS            HTG.
SOURCE              Mus musculus (house mouse)
ORGANISM            Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 213144)
Li, S., Tian, R. and Roe, B.A.
Mus musculus BAC Clone rp23-387f2
Unpublished
2 (bases 1 to 213144)
Li, S., Tian, R. and Roe, B.A.
Direct Submission
Submitted (10-JUL-2001) Department of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
3 (bases 1 to 213144)
Li, S., Tian, R. and Roe, B.A.
Direct Submission
Submitted (30-MAR-2002) Department of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
4 (bases 1 to 213144)
Li, S., Tian, R. and Roe, B.A.
Direct Submission
Submitted (02-APR-2002) Department of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
5 (bases 1 to 213144)
Li, S., Tian, R. and Roe, B.A.
Direct Submission
Submitted (16-JUL-2002) Department of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
6 (bases 1 to 213144)
Li, S., Tian, R. and Roe, B.A.
Direct Submission
Submitted (15-NOV-2002) Department of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
COMMENT On Nov 15, 2002 this sequence version replaced gi:21844616.
----- Genome Center
Center: Department Of Chemistry And Biochemistry
The University Of Oklahoma
Center code:UOKNOR

FEATURES             Location/Qualifiers
     source            1..213144
                        /organism="Mus musculus"
                        /mol_type="genomic DNA"
                        /strain="C57BL/6J"
                        /db_xref="taxon:10090"
                        /chromosome="17"
                        /clone="rp23-387f2"
                        /clone_lib="RPCI - 23 Female (C57BL/6J) Mouse BAC Library"

ORIGIN
Query Match      100.0%; Score 25; DB 9; Length 213144;
Best Local Similarity 100.0%; Pred. No. 0.37;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAAAACTGGAATCTCAGGCTGAGA 25
    |||||
Db 30614 AAAAACTGGAATCTCAGGCTGAGA 30590

```


AUTHORS

Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,
 Barna,N., Bastien,V., Boguslavsky,L., Bouckgalter,B., Brown,A.,
 Camarata,J., Campopiano,A., Choepel,Y., Colangelo,M., Collins,S.,
 Collymore,A., Cooke,P., DeRellano,K., Dewar,K., Diaz,J.S.,
 Dodge,S., Faro,S., Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J.,
 Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N.,
 Hagoos,B., Heaford,A., Horton,A., Hulme,W., Iliev,I., Johnson,R.,
 Jones,C., Karatas,A., LaRoque,K., Lamazares,R., Landers,T.,
 Lehoczy,J., Levine,R., Liu,G., MacLean,C., Macdonald,P.,
 Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K.,
 McPheeters,R., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
 Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H.,
 O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,
 Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R.,
 Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M.,
 Roy,A., Santos,R., Schauer,S., Schuback,R., Seaman,S., Severy,P.,
 Sounez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
 Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
 Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A.,
 Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J.,
 Zemek,L., Zimmer,A. and Zody,M.

TITLE

Direct Submission
 Submitted (13-DEC-2000) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA

JOURNAL

COMMENT

All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L11599

Center clone name: 58_O3

* NOTE: This record contains 32 individual
 * sequencing reads that have not been assembled into
 * contigs. Runs of N are used to separate the reads
 * and the order in which they appear is completely
 * arbitrary. Low-pass sequence sampling is useful for
 * identifying clones that may be gene-rich and allows
 * overlap relationships among clones to be deduced.
 * However, it should not be assumed that this clone
 * will be sequenced to completion. In the event that
 * the record is updated, the accession number will
 * be preserved.

* 1 692: contig of 692 bp in length
 * 693 792: gap of 100 bp
 * 793 1495: contig of 703 bp in length
 * 1496 1595: gap of 100 bp
 * 1596 2282: contig of 687 bp in length
 * 2283 2382: gap of 100 bp
 * 2383 3095: contig of 713 bp in length
 * 3096 3195: gap of 100 bp
 * 3196 3896: contig of 701 bp in length
 * 3897 3996: gap of 100 bp
 * 3997 4695: contig of 699 bp in length
 * 4696 4795: gap of 100 bp
 * 4796 5226: contig of 731 bp in length
 * 5227 5628: gap of 100 bp
 * 5629 6340: contig of 713 bp in length
 * 6341 6440: gap of 100 bp
 * 6441 7142: contig of 703 bp in length
 * 7143 7242: gap of 100 bp
 * 7243 7933: contig of 691 bp in length
 * 7934 8033: gap of 100 bp
 * 8034 8749: contig of 716 bp in length
 * 8750 8849: gap of 100 bp
 * 8850 9560: contig of 711 bp in length
 * 9561 9660: gap of 100 bp
 * 9661 10365: contig of 705 bp in length
 * 10366 10465: gap of 100 bp

* 10466 11186: contig of 721 bp in length
 * 11187 11286: gap of 100 bp
 * 11287 11999: contig of 713 bp in length
 * 12000 12099: gap of 100 bp
 * 12100 12809: contig of 710 bp in length
 * 12810 12909: gap of 100 bp
 * 12910 13608: contig of 699 bp in length
 * 13609 13708: gap of 100 bp
 * 13709 14404: contig of 696 bp in length
 * 14405 14504: gap of 100 bp
 * 14505 15224: contig of 720 bp in length
 * 15225 15324: gap of 100 bp
 * 15325 16048: contig of 724 bp in length
 * 16049 16148: gap of 100 bp
 * 16149 16864: contig of 716 bp in length
 * 16865 16964: gap of 100 bp
 * 16965 17677: contig of 713 bp in length
 * 17678 17777: gap of 100 bp
 * 17778 18494: contig of 717 bp in length
 * 18495 18594: gap of 100 bp
 * 18595 19287: contig of 693 bp in length
 * 19288 19387: gap of 100 bp
 * 19388 20085: contig of 698 bp in length
 * 20086 20185: gap of 100 bp
 * 20186 20933: contig of 748 bp in length
 * 20934 21033: gap of 100 bp
 * 21034 21782: contig of 749 bp in length
 * 21783 21882: gap of 100 bp
 * 21883 22592: contig of 710 bp in length
 * 22593 22692: gap of 100 bp
 * 22693 23421: contig of 729 bp in length
 * 23422 23521: gap of 100 bp
 * 23522 24215: contig of 694 bp in length
 * 24216 24315: gap of 100 bp
 * 24316 25019: contig of 704 bp in length
 * 25020 25119: gap of 100 bp
 * 25120 25837: contig of 718 bp in length.

Location/Qualifiers

1..25837
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="8"
 /map="8"
 /clone="RP11-5803"
 /clone_lib="RPC1-11 Human Male BAC"
 693..752
 /estimated_length=100
 1496..1595
 /estimated_length=100
 2283..2382
 /estimated_length=100
 3096..3195
 /estimated_length=100
 3897..3996
 /estimated_length=100
 4696..4795
 /estimated_length=100
 5527..5626
 /estimated_length=100
 6340..6439
 /estimated_length=100
 7143..7242
 /estimated_length=100
 7934..8033
 /estimated_length=100
 8750..8849
 /estimated_length=100
 9561..9660
 /estimated_length=100
 10366..10465
 /estimated_length=100
 11187..11286

FEATURES

source

gap

gap

gap

gap

gap

gap

gap

gap

gap

gap

gap

gap

gap

gap /estimated_length=100
 gap 12000..12099
 gap /estimated_length=100
 gap 12810..12909
 gap /estimated_length=100
 gap 13609..13708
 gap /estimated_length=100
 gap 14405..14504
 gap /estimated_length=100
 gap 15225..15324
 gap /estimated_length=100
 gap 16049..16148
 gap /estimated_length=100
 gap 16865..16964
 gap /estimated_length=100
 gap 17678..17777
 gap /estimated_length=100
 gap 18495..18594
 gap /estimated_length=100
 gap 19288..19387
 gap /estimated_length=100
 gap 20086..20185
 gap /estimated_length=100
 gap 20934..21033
 gap /estimated_length=100
 gap 21783..21882
 gap /estimated_length=100
 gap 22593..22692
 gap /estimated_length=100
 gap 23422..23521
 gap /estimated_length=100
 gap 24216..24315
 gap /estimated_length=100
 gap 25020..25119
 gap /estimated_length=100

ORIGIN

Query Match 80.8%; Score 20.2; DB 14; Length 25837;

Best Local Similarity 88.0%; Pred. No. 95; Mismatches 3; Indels 0; Gaps 0;

Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAACTGGAATCTCAGGCTGAGA 25
 |||||
 Db 2847 AGAAACTGGAATCTCAGGCTGAAA 2823

RESULT 8

AL596124/c
 LOCUS AL596124 30005 bp DNA linear PRI 18-MAY-2005
 DEFINITION Human DNA sequence from clone RP11-141H1 on chromosome 10, complete sequence.

ACCESSION AL596124.6 GI:15131373

VERSION HTG.

KEYWORDS Homo sapiens (human)

SOURCE Homo sapiens

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

REFERENCE 1 (bases 1 to 30005)

AUTHORS Clark.S.

TITLE Direct Submission

JOURNAL Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,

Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vegas@sanger.ac.uk

Clone requests: clonerequest@sanger.ac.uk

On Aug 9, 2001 this sequence version replaced gi:15041967.

The following abbreviations are used to associate primary accession

numbers given in the feature table with their source databases:

Em., EMBL; Sw., SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information

on the WORMPEP database can be found at

http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence

was generated from part of bacterial clone contigs of human

chromosome 10, constructed by the Sanger Centre Chromosome 10

Mapping Group. Further information can be found at
 http://www.sanger.ac.uk/HGP/Chr10

RP11-141H1 is from the library RPCI-11.1 constructed by the group
 of Pieter de Jong. For further details see

http://www.chori.org/bacpac/home.htm

VECTOR: pBAC3.6

----- Genome Center

Center: Wellcome Trust Sanger Institute

Center code: SC

Web site: http://www.sanger.ac.uk

Contact: vegas@sanger.ac.uk

This sequence was finished as follows unless otherwise noted: all
 regions were either double-stranded or sequenced with an alternate
 chemistry or covered by high quality data (i.e., phred quality >=
 30); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by at least
 one subclone; and the assembly was confirmed by restriction digest,
 except on the rare occasion of the clone being a YAC.

FEATURES

source

1..30005

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="10"

/clone="RP11-141H1"

/clone_lib="RPCI-11.1"

2000

/note="Clone_right_end: RP11-347I22"

28006

/note="Clone_left_end: RP11-730A19"

of Pieter de Jong. For further details see
<http://www.chori.org/bacpac/home.htm>
 VECTOR: pBACe3.6
 ----- Genome Center
 Center: Wellcome Trust Sanger Institute
 Center code: SC
 Web site: <http://www.sanger.ac.uk>
 Contact: vega@sanger.ac.uk

This sequence was finished as follows unless otherwise noted: all
 regions were either double-stranded or sequenced with an alternate
 chemistry or covered by high quality data (i.e., phred quality >=
 30); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by at least
 one subclone; and the assembly was confirmed by restriction digest,
 except on the rare occasion of the clone being a YAC.

FEATURES

```

source
1..180569
  /organism="Homo sapiens"
  /mol_type="genomic DNA"
  /db_xref="taxon:9606"
  /chromosome="1"
  /clone_lib="RPCI-11.1"
  /note="Clone left_end: RP11-82L20"
  /locus_tag="RP11-82L20"
  /product="neural growth regulator 1"
  /protein_id="CAI16870.1"
  /db_xref="GI:55958413"
  /db_xref="InterPro:IPR003598"
  /db_xref="InterPro:IPR003599"
  /db_xref="InterPro:IPR007110"
  /db_xref="UniProt/TREMBL:Q5VT21"
  /translation="MDMMLLVGACSNQWLAVALLSCLLPSPCLPAGQSVDFPNAALQI
  QNVVDVDDGPTSCVQQTPTPTMQVHLTVQVPPKIYDISNDMTVNEGTVTLTCLATGK
  PPSISMVNMVRKGDATVLRCYLEDGASGLNLSIIFAGGDKSWDPKVSISTLNKRDYS
  LATGKPEPSISWRHISPSAKPFENGQYLDIYIGTRDQAGEYECSAENDVSFDPVRKVK
  VVNFAPTIQELKSGTVTPGRSLIRCEGAGVPPPAFVYKGEKLFNGQOQGIIONF
  LRSILTVNTVQEHGNTVCVAANKLGTNASLPLNPSTAGYGITGSDVLFSCWY
  YLKNATLQ"
  /join(complement(136081..136201),complement(117872..118023),
  complement(136081..136201),complement(117872..118023))
  /gene="NEGR1"
  /locus_tag="RP11-316C12.1-001"
  /standard_name="OTTHUMP00000011267"
  /note="match: proteins: Tr:CAD97961 Tr:Q8N440 Tr:Q8NAQ3"
  /codon_start=1
  /product="neural growth regulator 1"
  /protein_id="CAI16869.1"
  /db_xref="GI:55958413"
  /db_xref="InterPro:IPR003598"
  /db_xref="InterPro:IPR003599"
  /db_xref="InterPro:IPR007110"
  /db_xref="UniProt/TREMBL:Q5VT21"
  /translation="MDMMLLVGACSNQWLAVALLSCLLPSPCLPAGQSVDFPNAALQI
  QNVVDVDDGPTSCVQQTPTPTMQVHLTVQVPPKIYDISNDMTVNEGTVTLTCLATGK
  PPSISMVNMVRKGDATVLRCYLEDGASGLNLSIIFAGGDKSWDPKVSISTLNKRDYS
  LATGKPEPSISWRHISPSAKPFENGQYLDIYIGTRDQAGEYECSAENDVSFDPVRKVK
  VVNFAPTIQELKSGTVTPGRSLIRCEGAGVPPPAFVYKGEKLFNGQOQGIIONF
  LRSILTVNTVQEHGNTVCVAANKLGTNASLPLNPSTAGYGITGSDVLFSCWY
  YLKNATLQ"
  /join(complement(136081..136201),complement(117872..118023),
  complement(136081..136201),complement(117872..118023))
  /gene="NEGR1"
  /locus_tag="RP11-316C12.1-002"
  /standard_name="OTTHUMP00000011268"
  /codon_start=1
  /product="neural growth regulator 1"
  /protein_id="CAI16870.1"
  /db_xref="GI:55958414"
  /db_xref="InterPro:IPR003598"
  /db_xref="InterPro:IPR007110"
  /db_xref="UniProt/TREMBL:Q8N440"
  /translation="MQVHLTVQVPPKIYDISNDMTVNEGTVTLTCLATGKPPSISM
  RHISPSAKPFENGQYLDIYIGTRDQAGEYECSAENDVSFDPVRKVKVNVNFAPTIQEI
  KSGTVTPGRSLIRCEGAGVPPPAFVYKGEKLFNGQOQGIIONFSLILTVNTV
  QSHFGNTVCVAANKLGTNASLPLNPSTAGYGITGSDVLFSCWYLVLTSLTSFTSIF
  YLKNATLQ"
  /misc_feature
  180569
  /note="Clone right_end: RP11-82L20"

ORIGIN
Query Match      80.8%; Score 20.2; DB 8; Length 180569;
Best Local Similarity 88.0%; Pred. No. 84;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 AAAAACTGGAAATCTCAGCTGAGA 25
      |||||
Db      14388 AGAAACTGGAATCTCAGCTGAAA 14364

RESULT 10
AC156206/c
LOCUS      AC156206
DEFINITION Bos taurus clone CH240-39N6, *** SEQUENCING IN PROGRESS ***, 10
unordered pieces.
ACCESSION AC156206
VERSION   AC156206.2
KEYWORDS  HTG; HTGS PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE    Bos taurus (cow)
ORGANISM  Bos taurus

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.

1 (bases 1 to 193379)

Muzny, D. Marie, Metzker, M. Lee, Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D., Anyalebechi, V., Ayodeji, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, O., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buha, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Kocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, P., Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, M., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowals, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorushtewa, L., Loulseged, H., Lozada, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhney, S., McLeod, M. P., McNeill, T. Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwaokeme, O., Okwuonu, G., Olarnpungoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfankoch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L., L., Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, W., Saverly, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartabeyn, A., Sisson, I., Sitter, C. D., Smajic, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steidle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Umani, K., Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczek, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, J., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstock, G. and Gibbs, R. A.

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

gap

gap

gap

gap

gap

gap

gap

gap

individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

Center: Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

Project Information

Center project name: FCSW

Center clone name: CH240-39N6

Summary Statistics

Assembly program: Atlas 3.0;

Consensus quality: 186284 bases at least Q40

Consensus quality: 187266 bases at least Q30

Consensus quality: 188115 bases at least Q20

Estimated insert size: 187827; sum-of-contigs estimation

Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

NOTE: Estimated insert size may differ from sequence length

(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).

NOTE: This is a 'working draft' sequence. It currently

consists of 10 contigs. The true order of the pieces

is not known and their order in this sequence record is

arbitrary. Gaps between the contigs are represented as

runs of N, but the exact sizes of the gaps are unknown.

This record will be updated with the finished sequence

as soon as it is available and the accession number will

be preserved.

1 102587: contig of 102587 bp in length

102588 102637: gap of 50 bp

102638 123249: contig of 20612 bp in length

123250 124403: gap of 1154 bp

124404 126315: contig of 1912 bp in length

126316 127559: gap of 1244 bp

127560 128794: contig of 1235 bp in length

128795 129875: gap of 1081 bp

129876 159012: contig of 29137 bp in length

159013 159801: gap of 789 bp

159802 168291: contig of 8490 bp in length

168292 168341: gap of 50 bp

168342 174113: contig of 5772 bp in length

174114 174163: gap of 50 bp

174164 191240: contig of 16977 bp in length

191241 192243: contig of 1003 bp in length

192244 192343: gap of unknown length

192344 193379: contig of 1036 bp in length.

Location/Qualifiers

1. 193379

/organism="Bos taurus"

/mol_type="genomic DNA"

/db_xref="taxon:9913"

/clone="CH240-39N6"

102588. 102637

/estimated length=50

123250. 124403

/estimated length=1154

126316. 127559

/estimated length=1244

128795. 129875

/estimated length=1081

159013. 159801

/estimated length=789

168292. 168341

/estimated length=50

174114. 174163

/estimated length=50

191141. 191240

```

/estimated_length=unknown
192244..192343
/estimated_length=unknown

ORIGIN
Query Match      80.8%; Score 20.2; DB 14; Length 193379;
Best Local Similarity 88.0%; Pred. No. 83;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAACTGGAATCTCAGGCTGAGA 25
|||||
DB 54135 AAAAACTGGAACCTCAAACTGAGA 54111

RESULT 11
AC053464/c
LOCUS          194019 bp      DNA      linear      HTG 30-AUG-2001
DEFINITION    Homo sapiens chromosome 10 clone RP11-383B5, WORKING DRAFT
SEQUENCE. 12 unordered pieces.
ACCESSION     AC053464
VERSION       HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_CANCELLED.
KEYWORDS      Homo sapiens (human)
SOURCE        Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
              Homnidae; Homo.
REFERENCE     Smith,D.R.
              Genome Therapeutics Corporation Sequencing Center: Human Genome
              Sequence Data
              Unpublished
              2 (bases 1 to 194019)
JOURNAL
REFERENCE     Smith,D.R.
              Direct Submission
              Submitted (16-APR-2000) Genome Therapeutics Corporation, 100 Beaver
              Street, Waltham, MA 02453, USA
              On Apr 25, 2001 this sequence version replaced gi:10944437.
COMMENT      -----
              Genome Center
              Center: Genome Therapeutics Corporation
              Center code: GTC
              Web site: http://www.genomecorp.com/
              Contact: gtc-seqcenter@genomecorp.com
              -----
              Project Information
              -----
              Center project name: hg209
              -----
              Summary Statistics
              -----
              Sequencing vector: N/A
              Chemistry: Dye-terminator Big Dye; 100% of reads
              Assembly program: Phrap; version 990315
              Consensus quality: 189621 bases at least Q40
              Consensus quality: 191316 bases at least Q30
              Consensus quality: 192144 bases at least Q20
              Insert size: 196098; sum-of-contigs
              Quality coverage: 6.1x in Q20 bases; sum-of-contigs
              -----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 12 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 1409: contig of 1409 bp in length
* 1509: gap of unknown length
* 3294: contig of 1785 bp in length
* 3394: gap of unknown length
* 3395: contig of 2024 bp in length
* 5418: gap of unknown length
* 5519: contig of 1048 bp in length
* 6566: gap of unknown length
* 6667: contig of 3193 bp in length
* 9860: gap of unknown length
* 9860: gap of unknown length

us-10-719-900-30.rge
```

```

* 9960 15190: contig of 5231 bp in length
* 15191 15290: gap of unknown length
* 15291 25306: contig of 10016 bp in length
* 25307 25406: gap of unknown length
* 25407 39785: contig of 14379 bp in length
* 39786 39885: gap of unknown length
* 39886 57462: contig of 17577 bp in length
* 57463 57562: gap of unknown length
* 57563 79693: contig of 22131 bp in length
* 79694 79793: gap of unknown length
* 79794 122822: contig of 43029 bp in length
* 122823 122923: gap of unknown length
* 122923 194019: contig of 71097 bp in length.

FEATURES
Location/Qualifiers
1..194019
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="10"
/clone_lib="RPC1-11"
/clone="RP11-383B5"
1..1409
/feature="assembly_name:Contig7"
1410..1509
/feature="estimated_length=unknown"
1510..3294
/feature="assembly_name:Contig12"
3295..3394
/feature="estimated_length=unknown"
3395..5418
/feature="assembly_name:Contig13"
5419..5518
/feature="estimated_length=unknown"
5519..6566
/feature="assembly_name:Contig14"
6567..6666
/feature="estimated_length=unknown"
6667..9859
/feature="assembly_name:Contig16"
9860..9959
/feature="estimated_length=unknown"
9960..15190
/feature="assembly_name:Contig17"
15191..15290
/feature="estimated_length=unknown"
15291..25306
/feature="assembly_name:Contig18"
25307..25406
/feature="estimated_length=unknown"
25407..39785
/feature="assembly_name:Contig19"
39786..39885
/feature="estimated_length=unknown"
39886..57462
/feature="assembly_name:Contig20"
57463..57562
/feature="estimated_length=unknown"
57563..79693
/feature="assembly_name:Contig21"
79694..79793
/feature="estimated_length=unknown"
79794..122822
/feature="assembly_name:Contig22"
122823..122922
/feature="estimated_length=unknown"
122923..194019
/feature="assembly_name:Contig23"

ORIGIN
Query Match      80.8%; Score 20.2; DB 14; Length 194019;
Best Local Similarity 88.0%; Pred. No. 83;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 1 AAAAACTGGAATCTCAGGCTGAGA 25
    ||| ||||| ||||| ||||| |||||
Db 94166 AATGACTGGAATCTCAGGCAGAGA 94142

RESULT 12
AC013446
LOCUS AC013446 208202 bp DNA linear HTG 07-JUL-2000
DEFINITION Homo sapiens chromosome 10 clone RP11-513P21, WORKING DRAFT
ACCESSION AC013446
VERSION AC013446.3 GI:7923997
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE 1 (bases 1 to 208202)
AUTHORS Waterston,R.H.
TITLE The sequence of Homo sapiens clone
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 208202)
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (11-NOV-1999) Genome Sequencing Center, Washington
University School of Medicine, 444 Forest Park Parkway, St. Louis,
MO 63108, USA
COMMENT On May 18, 2000 this sequence version replaced gi:6850545.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H.NH0513P21
----- Summary Statistics -----
Sequencing vector: M13; 84%
Sequencing vector: plasmid; 16%
Chemistry: Dye-primer ET; 84% of reads
Chemistry: Dye-terminator Big Dye; 16% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 196339 bases at least Q40
Consensus quality: 199886 bases at least Q30
Consensus quality: 202038 bases at least Q20
Insert size: 215000; agarose-fp
Insert size: 206002; sum-of-contigs
Quality coverage: 4.38 in Q20 bases; agarose-fp
Quality coverage: 4.26 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 23 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
* as soon as it is available and the accession number will
* be preserved.
*
* 1 1699: contig of 1699 bp in length
* 1700: gap of unknown length
* 1800: contig of 1441 bp in length
* 3241: gap of unknown length
* 3341: contig of 2875 bp in length
* 6216: gap of unknown length
* 6316: contig of 2740 bp in length
* 9056: gap of unknown length
* 9156: contig of 2689 bp in length
* 11845: gap of unknown length
* 11945: contig of 3647 bp in length
* 15591: gap of unknown length
* 15692: contig of 2566 bp in length
* 18258: gap of unknown length
* 18358: contig of 5081 bp in length
* 23438: contig of 1699 bp in length
* 23539: gap of unknown length
* 27554: gap of unknown length
* 32489: gap of unknown length
* 34085: contig of 3817 bp in length
* 36406: gap of unknown length
* 36505: gap of unknown length
* 40594: gap of unknown length
* 40694: contig of 7089 bp in length
* 47783: gap of unknown length
* 47883: contig of 7302 bp in length
* 55185: gap of unknown length
* 55285: contig of 8044 bp in length
* 63329: gap of unknown length
* 63429: contig of 10673 bp in length
* 74102: gap of unknown length
* 74202: contig of 10849 bp in length
* 85051: gap of unknown length
* 85151: contig of 15419 bp in length
* 100570: gap of unknown length
* 100670: contig of 15567 bp in length
* 116237: gap of unknown length
* 116337: contig of 15312 bp in length
* 131649: gap of unknown length
* 131749: contig of 17078 bp in length
* 148827: gap of unknown length
* 148927: contig of 28223 bp in length
* 175150: gap of unknown length
* 175250: contig of 32953 bp in length.

FEATURES
Location/Qualifiers
1..208202
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="10"
/clone="RP11-513P21"
misc_feature 1..1699
/note="assembly_name:Contig10"
gap 1700..1799
/estimated_length=unknown
misc_feature 1800..3240
/note="assembly_name:Contig11"
gap 3241..3340
/estimated_length=unknown
misc_feature 3341..6215
/note="assembly_name:Contig12"
gap 6216..6315
/estimated_length=unknown
misc_feature 6316..9055
/note="assembly_name:Contig13"
gap 9056..9155
/estimated_length=unknown
misc_feature 9156..11844
/note="assembly_name:Contig14"
gap 11845..11944
/estimated_length=unknown
misc_feature 11945..15591
/note="assembly_name:Contig15"
gap 15592..15691
/estimated_length=unknown
misc_feature 15692..18257
/note="assembly_name:Contig16"
gap 18258..18357
/estimated_length=unknown
misc_feature 18358..23438
/note="assembly_name:Contig17"
gap 23439..23538
/estimated_length=unknown
misc_feature 23539..27553
/note="assembly_name:Contig18"
gap 27554..27653
/estimated_length=unknown
```


COMMENT

On Jun 28, 2005 this sequence version replaced gi:58037610.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: FAXI
Center clone name: CH240-17N21
----- Summary Statistics

Assembly program: Atlas 3.0;
Consensus quality: 224136 bases at least Q40
Consensus quality: 225748 bases at least Q30
Consensus quality: 227236 bases at least Q20
Estimated insert size: 226781; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 20 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1
1302: contig of 1302 bp in length
1303: 2070: gap of 768 bp
2071: 4765: contig of 2695 bp in length
4766: 4815: gap of 50 bp
4816: 22885: contig of 18070 bp in length
22886: 22935: gap of 50 bp
22936: 36075: contig of 13140 bp in length
36076: 36125: gap of 50 bp
36126: 69777: contig of 33652 bp in length
69778: 69827: gap of 50 bp
69828: 97935: contig of 28108 bp in length
97936: 97985: gap of 50 bp
97986: 101342: contig of 3357 bp in length
101343: 101392: gap of 50 bp
101393: 105292: contig of 3900 bp in length
105293: 105342: gap of 50 bp
105343: 137936: contig of 32594 bp in length
137937: 137986: gap of 50 bp
137987: 143769: contig of 5783 bp in length
143770: 143819: gap of 50 bp
143820: 145030: contig of 1211 bp in length
145031: 145080: gap of 50 bp
145081: 150012: contig of 4932 bp in length
150013: 150062: gap of 50 bp
150063: 154610: contig of 4548 bp in length
154611: 154660: gap of 50 bp
154661: 165400: contig of 10740 bp in length
165401: 165450: gap of 50 bp
165451: 179172: contig of 13722 bp in length
179173: 179222: gap of 50 bp
179223: 224914: contig of 45692 bp in length
224915: 224964: gap of 50 bp
224965: 226156: contig of 1192 bp in length
226157: 226256: gap of unknown length

* 226257 227672: contig of 1416 bp in length
* 227673 227772: gap of unknown length
* 227773 228777: contig of 1005 bp in length
* 228778 228877: gap of unknown length
* 228878 231342: contig of 2465 bp in length.

FEATURES
source

Location/Qualifiers
1..231342
/organism="Bos taurus"
/mol_type="genomic DNA"
/db_xref="taxon:9913"
/clone="CH240-17N21"
1303..2070
/estimated_length=768
4766..4815
/estimated_length=50
22886..22935
/estimated_length=50
36076..36125
/estimated_length=50
69778..69827
/estimated_length=50
97936..97985
/estimated_length=50
101343..101392
/estimated_length=50
105293..105342
/estimated_length=50
137937..137986
/estimated_length=50
143770..143819
/estimated_length=50
145031..145080
/estimated_length=50
150013..150062
/estimated_length=50
154611..154660
/estimated_length=50
165401..165450
/estimated_length=50
179173..179222
/estimated_length=50
224915..224964
/estimated_length=50
226157..226256
/estimated_length=unknown
227673..227772
/estimated_length=unknown
228778..228877
/estimated_length=unknown

ORIGIN

Query Match 80.8%; Score 20.2; DB 14; Length 231342;
Best Local Similarity 88.0%; Pred. No. 82;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAACTGGAATCTCAGGCTGAGA 25
|||||

Db 93857 AAAAACTGGAACCTCATGCTGAGA 93833
|||||

RESULT 14

BD212075
LOCUS BD212075 300 bp DNA linear PAT 17-JUL-2003
DEFINITION Novel human genes and gene expression products ii.
ACCESSION BD212075
VERSION BD212075.1 GI:33021845
KEYWORDS JP 2002519000-A/217.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 300)

```

AUTHORS Williams,L.T., Escobedo,J., Innis,M.A., Garcia,P.D., Klinger,J.S.,
Reinhard,C., Giese,K., Randazzo,F., Kennedy,G.C., Pot,D.,
Kassam,A., Lamsom,G., Drmanac,R., Crkvenjakov,R., Dickson,M.,
Drmanac,S., Labat,I., Leshkowitz,D., Kita,D., Garcia,V., Jones,L.W.
and Crain,B.S.
TITLE Novel human genes and gene expression products ii
JOURNAL Patent: JP 2002519000-A 217 02-JUL-2002;
COMMENT CHIRON CORP.HYSEQ INC
OS Homo sapiens (human)
PN JP 2002519000-A/217
PD 02-JUL-2002
PF 28-JAN-1999 JP 2000556580
PR 28-JAN-1998 US 60/072910,24-FEB-1998 US 60/075954 PR
31-MAR-1998 US 60/080114,03-APR-1998 US 60/080515 PR
03-APR-1998 US 60/080666,21-OCT-1998 US 60/105234 PR
28-OCT-1998 US 60/105877
PI LOUIS T WILLIAMS,JAIME ESCOBEDO,MICHAEL A INNIS,PABLO PI
DOMINGUEZ GARCIA,
PI JULIE SUDDUTH KLINGER,CHRISTOPH REINHARD,KLAUSE GIESE,FILIPPO
PI RANDAZZO,
PI GIULIA C KENNEDY,DAVID POT,ALTAF KASSAM,GEORGE LAMSON,RADOJE
PI DRMANAC,
PI RADOMIR CRKVENJAKOV,MARK DICKSON,SNEZANA DRMANAC,IVAN LABAT,
PI DENA LESHKOWITZ,DAVID KITA,VERONICA GARCIA,LEE WILLIAM JONES,
PI BIRJIT STACHE CRAIN
PC C12N15/09,C12N15/09,C07K14/47,C07K14/82,C07K16/18,C12N1/15, PC
C12N1/19,
PC C12N1/21,C12N5/10,C12Q1/68,C12N15/00,C12N5/00,C12N15/00 CC
Novel human genes and gene expression products ii FH Key
Location/Qualifiers
FT source 1. .300
/organism="Homo sapiens (human)".
/locus Location/Qualifiers
1. .300
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 80.0%; Score 20; DB 6; Length 300;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4 AAACCTGGAATCTCAGGCTGA 23
|||||
Db 104 AAACCTGGAATCTCAGGCTGA 123
|||||

RESULT 15
BD215003 758 bp DNA linear PAT 17-JUL-2003
LOCUS Novel human genes and gene expression products ii.
DEFINITION BD215003
ACCESSION BD215003
VERSION BD215003.1 GI:33024773
KEYWORDS JP 2002519000-A/3145.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE 1 (bases 1 to 758)
AUTHORS Williams,L.T., Escobedo,J., Innis,M.A., Garcia,P.D., Klinger,J.S.,
Reinhard,C., Giese,K., Randazzo,F., Kennedy,G.C., Pot,D.,
Kassam,A., Lamsom,G., Drmanac,R., Crkvenjakov,R., Dickson,M.,
Drmanac,S., Labat,I., Leshkowitz,D., Kita,D., Garcia,V., Jones,L.W.
and Crain,B.S.
TITLE Novel human genes and gene expression products ii
JOURNAL Patent: JP 2002519000-A 3145 02-JUL-2002;
COMMENT CHIRON CORP.HYSEQ INC
OS Homo sapiens (human)
PN JP 2002519000-A/3145
PD 02-JUL-2002
PF 28-JAN-1999 JP 2000556580

AUTHORS Williams,L.T., Escobedo,J., Innis,M.A., Garcia,P.D., Klinger,J.S.,
Reinhard,C., Giese,K., Randazzo,F., Kennedy,G.C., Pot,D.,
Kassam,A., Lamsom,G., Drmanac,R., Crkvenjakov,R., Dickson,M.,
Drmanac,S., Labat,I., Leshkowitz,D., Kita,D., Garcia,V., Jones,L.W.
and Crain,B.S.
TITLE Novel human genes and gene expression products ii
JOURNAL Patent: JP 2002519000-A 3145 02-JUL-2002;
COMMENT CHIRON CORP.HYSEQ INC
OS Homo sapiens (human)
PN JP 2002519000-A/3145
PD 02-JUL-2002
PF 28-JAN-1999 JP 2000556580

PR 28-JAN-1998 US 60/072910,24-FEB-1998 US 60/075954 PR
31-MAR-1998 US 60/080114,03-APR-1998 US 60/080515 PR
03-APR-1998 US 60/080666,21-OCT-1998 US 60/105234 PR
28-OCT-1998 US 60/105877
PI LOUIS T WILLIAMS,JAIME ESCOBEDO,MICHAEL A INNIS,PABLO PI
DOMINGUEZ GARCIA,
PI JULIE SUDDUTH KLINGER,CHRISTOPH REINHARD,KLAUSE GIESE,FILIPPO
PI RANDAZZO,
PI GIULIA C KENNEDY,DAVID POT,ALTAF KASSAM,GEORGE LAMSON,RADOJE
PI DRMANAC,
PI RADOMIR CRKVENJAKOV,MARK DICKSON,SNEZANA DRMANAC,IVAN LABAT,
PI DENA LESHKOWITZ,DAVID KITA,VERONICA GARCIA,LEE WILLIAM JONES,
PI BIRJIT STACHE CRAIN
PC C12N15/09,C12N15/09,C07K14/47,C07K14/82,C07K16/18,C12N1/15, PC
C12N1/19,
PC C12N1/21,C12N5/10,C12Q1/68,C12N15/00,C12N5/00,C12N15/00 CC
=A,T,C or G
FH Key Location/Qualifiers
FT misc feature (1)..(758).
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
FEATURES
source
ORIGIN
Query Match 80.0%; Score 20; DB 6; Length 758;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4 AAACCTGGAATCTCAGGCTGA 23
|||||
Db 148 AAACCTGGAATCTCAGGCTGA 167
|||||

RESULT 16
ABI68750 2272 bp mRNA linear PRI 18-JUN-2005
LOCUS Macaca fascicularis testis cDNA, clone: QtsA-14569, similar to
human tubby like protein 4 (TULP4), mRNA, RefSeq: NM_020245.2.
ACCESSION ABI68750.1 GI:67969004
VERSION ABI68750.1 GI:67969004
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Macaca fascicularis (crab-eating macaque)
ORGANISM Macaca fascicularis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Cercopitheidae; Cercopithecinæ; Macaca.
REFERENCE 1
AUTHORS International consortium for macaque cDNA sequencing and analysis.
TITLE DNA sequences of macaque genes expressed in brain or testis and its
evolutionary implications
JOURNAL Unpublished
REFERENCE 2
AUTHORS Oeada,N., Hirata,M., Tanuma,R., Kusuda,J., Hida,M., Suzuki,Y.,
Sugano,S., Gojobori,T., Shen,J.C.-K., Wu,C.I. and Hashimoto,K.
TITLE Substitution rate and structural divergence of 5'UTR evolution:
Comparative analysis between human and cynomolgus monkey cDNAs
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 2272)
AUTHORS Hashimoto,K., Kusuda,J. and Sugano,S.
TITLE Direct Submission
JOURNAL Submitted (18-MAR-2004) Katsuyuki Hashimoto, National Institute of
Infectious Diseases, Division of Genetic Resources; 23-1, Toyama
1-chome, Shinjuku-ku, Tokyo, 162-8640, Japan
(E-mail:khashi@nih.go.jp, URL:http://www.nih.go.jp/yoken/genebank/,
Tel:81-3-5285-1111(ex.2120), Fax:81-3-5285-1181)
COMMENT The International consortium for macaque cDNA sequencing and
analysis consists of: Department of Virology and Human Genome
Center, Institute of Medical Science, The University of Tokyo,
Tokyo, Japan; Division of Genetic Resources, National Institute of
Infectious Diseases of Japan, Tokyo, Japan; National Health
Research Institute, Taipei, Taiwan; Institute of Molecular Biology,

```

Academia Sinica, Taipei, Taiwan; Department of Ecology & Evolution, University of Chicago, Chicago, IL, USA; Center for Information Biology; National Institute of Genetics of Japan, Mishima, Japan. Clone distribution: clone distribution information can be found at: <http://www.nih.go.jp/yoken/genebank/>

Lab host: TOP10

Vector: pME18S-FL3 (Acc.No. AB009864)

R. Site1: DraIII (CACTGTGTC)

R. Site2: DraIII (CACATGTC)

Description: 1st strand cDNA was primed with an oligo(dT) primer (ATGGCCCTTTTCTTTTCTTTT); double-stranded cDNA was synthesized using specific 5' and 3' primers and amplified by PCR. The PCR product was digested with SfiI and size selection was performed to exclude fragments <1.5kb. The SfiI-digested PCR product was cloned into distinct DraIII sites of pME18S-FL3. XhoI sites just outside the DraIII sites can be used to isolate the cDNA insert. Libraries were constructed by oligo-capping method. Libraries were made from:

OccB: cerebellum cortex

QnpA: parietal lobe

QtrA: temporal lobe right

QflA: frontal lobe left

QmoA: medulla oblongata

QbsA: brain stem

QorA: occipital lobe right

QtsA: testis

Custom primers were used for 5' and 3'-end sequencing. The full-insert sequencing was done by primer-walking method using ABI DNA sequencer.

FEATURES

Location/Qualifiers

1..2272

/organism="Macaca fascicularis"

/mol_type="mRNA"

/db_xref="taxon:9541"

/clone="QtsA-14569"

/sex="male"

/clone_lib="macaque cDNA library QtsA"

/dev_stage="adult"

148..603

/note="unnamed protein product; Homo sapiens tubby like protein 4 (TULP4), mRNA, RefSeq: NM_020245.2"

/codon_start=1

/protein_id="BAE00858.1"

/db_xref="GI:67968005"

/translation="MSDNDIDVESDADKRAHNALEKRRDHIXDSFSLRDSVPSL QGEKARAILDKATEYIQYMRKNHTQQDIDDLKRONALLEQVRALKARSSAQL QTNYPSSDNLNTAKGTSISAFDGSBSSSEPEPQSRKKLRMEAS"

ORIGIN

Query Match 80.0%; Score 20; DB 8; Length 2272;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 AAACCTGGAATCTCAGGCTGA 23

|||||

Db 1240 AAACCTGGAATCTCAGGCTGA 1259

RESULT 17

AB037818

LOCUS AB037818 4810 bp mRNA linear PRI 14-MAR-2000

DEFINITION Homo sapiens mRNA for KIAA1397 protein, partial cds.

AB037818

ACCESSION

AB037818.1 GI:7243174

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

REFERENCE

1 (sites)

1 Nagase, T., Kikuno, R., Iehikawa, K.I., Hirokawa, M. and Ohara, O.

Prediction of the coding sequences of unidentified human genes.

XVI. The complete sequences of 150 new cDNA clones from brain which

code for large proteins in vitro

DNA Res. 7 (1), 65-73 (2000)

10718198

2 (bases 1 to 4810)

CHARA.O., Nagase, T. and Kikuno, R.

Direct Submission

JOURNAL

Submitted (31-JAN-2000) Osamu Ohara, Kazusa DNA Research Institute,

Chara.O., Nagase, T. and Kikuno, R.

Laboratory of DNA Technology; 1532-3 Yana, Kisarazu, Chiba

292-0812, Japan (E-mail: cdnainfo@kazusa.or.jp.

URL: <http://www.kazusa.or.jp/huge/>, Tel: +81-438-52-3913,

Fax: +81-438-52-3914)

Location/Qualifiers

1..4810

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="hj08255"

/tissue type="brain"

/clone_lib="pBluescriptII SK plus"

1..4810

/gene="KIAA1397"

<1334..3394

/gene="KIAA1397"

/note="Start codon is not identified."

/codon_start=1

/product="KIAA1397 protein"

/protein_id="BAA92635.1"

/db_xref="GI:7243175"

/translation="VKQPIESMYAAVEHGVLCSDSNILCLSWKGRVPSKEKEKPV
RRYYEGWLATNGRGVGVVFTTSSHCRRDRSTPQRINFLRGNHSEVILVRWNEPY
QKLCADADGGIFVMIQVEGRWSVELVNDRGAQVSDFTWSDHGTQALISYRDGFLV
SVSGORHWSSEINLESOITCGIWTDPDQQLVFGTAGQVIVMDCHGRMLAHVLHESD
GVLGMSWNPPIFLVEDSSESDSDDYAPDQGPAAVPIPVQNIKPLLTVSTSGDIS
LMNNYDDLSPTVIRSGLVKVAQCTQGLDLAVAGMERQTLGELPNGLPKLSAMVKF
YNNRGELITLTLVQRPILISICWGRDSRLMASGPALYVVRVHSSQLLCCQA
IASTLRDDKVSGLTLPRLCSYLSTAFIPTIKPIPDNNMRDFVSPSAGNERLHC
TMKRTEDDPEVGGCYTLLEYLGLVPLIKGRRIKSLRPFVIMDPRDTSKPDIEYG
NSLSTVTDSCNSDSDIELSDWAASKSPKISRAKSPKLPRISIRAKSPKLPRAPRA
AQELSRSPRLRPSVGSPLTRPEPPEDITOHNVLAQVTSNIGTKFKIVGLAAP
LPTNLGAVIYKTSLLHLQPRQMTIYLPVPRKISMVYINLPVFNPNVFEDEDDLPDGA
VWTD"

ORIGIN

Query Match 80.0%; Score 20; DB 8; Length 4810;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 AAACCTGGAATCTCAGGCTGA 23

|||||

Db 964 AAACCTGGAATCTCAGGCTGA 983

RESULT 18

AX405807

LOCUS

AX405807

DEFINITION

AX405807

ACCESSION

AX405807.1

VERSION

GI:214339053

KEYWORDS

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

REFERENCE

1

AUTHORS

Xue, A.J., Yang, Y., Wehrman, T. and Drmanac, R.T.

Novel nucleic acids and polypeptides

Patent: WO 0222660-A 222 21-MAR-2002;

HYSEQ, INC. (US)

LOCATION/Qualifiers

1..5174

/organism="Homo sapiens"

```

/mol_type="unassigned DNA"
/db_xref="taxon:9606"
1358..3394
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAD34914.1"
/db_xref="GI:21439054"

CDS
/translation="MTAAVHGHPVLCSDSNILCSWKGRVPKSEKEKPCVRRRYVEEG
WLATNGRGVGVFTTSSHCRRDSTPQINFNLHGNSSEVLVRNNEYPQKLATCDA
DGGIPIWIOYGRWSVLVNDRAQVSDFTWGHSDTQALISYRDGVLVGVSGQGRH
SEINLSQITCGIWTDDQVGLPTGADQVIMDCIKRLHALLIHESDQVGLGMSN
YPIFLVEDSESDTDDVYAPDPGAAVPIPVNTIKPLLTVSTSGDLSLMNYYDL
SPVIRSGIKVEYVAOCTGDLAVAMERQITQELPNGLLKSAWFKVYNRGHI
FLTDTUVRPTISICWGRDSRLLMASGPALYVVRVHRVYSSILQLCQAALSTLED
KDVSKUTLPRLCSYLSTAFIPITKPIIPDNNMRDVFSPSAGNERLHCQATLRDD
PEVGGCYTLIYLELGLVPIIKGRRI SKLRPEFVIMDRPTDSKPDIEIYVNSLISVI
DSCGSDSDSIELSDWAAKSPKISRASKSPKPLRISIEARKSPKPLRAAQELSRP
RULRKPSPKSPKLRPEFPFEDIIOHNYLAQVTSNIGWTKFKIIVGLAAFLPTNLGAV
LYKTSI.LHQPROMTITLVPRVKITSMDYNI.LRVFNNSVFESEDDLLPGDAVWT"

```

```

ORIGIN
      80.0%;   Score 20;   DB 6;   Length 5174;
Query Match      Best Local Similarity 100.0%;   Pred. No. 1.3e+02;
Matches 20;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;

QY      4   AAACCTGGAAATCTCAGGCTGA 23
          |||||
Db      964   AAACCTGGAAATCTCAGGCTGA 983
          |||||

```

RESULT	19
AF288480	
LOCUS	5183 bp mRNA linear PRI 20-AUG-2000
DEFINITION	Homo sapiens tubby super-family protein (TUSP) mRNA, complete cds, alternatively spliced.
ACCESSION	AF288480
VERSION	AF288480.1 GI:9859153
KEYWORDS	.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE	1. (bases 1 to 5183) Li,Q.-Z., Wang,C. and She,J.-X. Molecular cloning of TUSP gene, a novel member of tubby gene super-family, its expression, alternative splicing and chromosome localization on 6q25-26, from human and mouse
AUTHORS	Unpublished
TITLE	
JOURNAL	
REFERENCE	2 (bases 1 to 5183) Li,Q.-Z., Wang,C., Shi,J.-D., Cruze,P. and She,J.-X. Direct Submission Submitted (19-JUL-2000) Pathology, Immunology and Laboratory Science, University of Florida, 1600 SW Archer Road, Rm. D6-15, Gainesville, FL 32610. USA
AUTHORS	
TITLE	
JOURNAL	

```

FEATURES
source
Location/Qualifiers
1. .5183
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="6"
/map="6q25-q26"
1. .5183
/gene="TUSP"
1358. .3394
/gene="TUSP"
/note="alternatively spliced"
/codon_start=1
/product="tubby super-family protein"
/protein_id="AAG01020.1"
/db_xref="GI:9858154"
/translation="MYAAVHEGPGVLCSDSNILCLSWKGRVPKSEKEKPVCRYYRREG
WLATNGRGVVGVTFSTSHCRDRRSTPQRIINFLRHNSEVILVRMNPYQKLATCD"

```

	Query Match	80.0%;	Score 20;	DB 8;	Length 5183;
Best Local Similarity	100.0%;	Pred. No. 1.3e+02;			
Matches 20;	Conservative	0;	Mismatches	0;	Indels 0; Gaps 0;
QY	4	AAACTGGAAATCTCAGGCTGA	23		
Db	964	AAACTGGAAATCTCAGGCTGA	983		

RESULT	20
LOCUS	CS033671
DEFINITION	Sequence 3177 from Patent WO2005016962.
ACCESSION	CS033671
VERSION	CS033671.1
KEYWORDS	. GI:60732618
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
	linear
	PAT 10-MAR-2005

REFERENCE

AUTHORS Abbas,A., Clark,H., Ouyang,W., Williams,M.P., Wood,W.I. and Wu,T.D.

TITLE Compositions and methods for the treatment of immune related diseases

JOURNAL Patent: WO 2005016962-A 3177 24-FEB-2005;

FEATURES Genentech, Inc. (US)

source Location/Qualifiers

1..11127

/organism="Homo sapiens"

/mol_type="unassigned DNA"

/db_xref="taxon:9606"

ORIGIN

Query Match	80.0%;	Score 20;	DB 6;	Length 11127;
Best Local Similarity	100.0%;	Pred. No. 1.2e+02;		
Matches 20;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	4	AAACTGGAATCTCAGGCTGA	23	
Db	964	AAACTGGAATCTCAGGCTGA	983	
RESULT 21				
CS042623				
LOCUS	CS042623	11127 bp	DNA	linear
DEFINITION	Sequence 3177 from Patent WO2005019258.			
ACCESSION	CS042623			
VERSION	CS042623.1	GI:61849791		
KEYWORDS	.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE	1			
AUTHORS	Abbas, A., Clark, H., Ouyang, W., Williams, P. M., Wood, W. I. and Wu, T. D.			
TITLE	Compositions and methods for the treatment of immune related diseases			
JOURNAL	Patent: WO 2005019258-A 3177 03-MAR-2005;			
	Genentech, Inc. (US)			

as compressions and repeats; all regions were covered by at least one subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

FEATURES

source

Location/Qualifiers

```
1. .180635
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="6"
/clone="RP11-732M18"
/clone_lib="RPCI-11.3"
```

misc_feature

```
1
/notes="Clone left end: RP11-732M18"
join(complement(162776..164384),complement(63851..63979),
complement(47147..47308),complement(27868..28048),
complement(24776..24910),complement(15315..15481),
complement(AL353800.10:33419..33643),
complement(AL353800.10:32105..32339),
complement(AL353800.10:23662..23806),
complement(AL353800.10:19677..19821),
complement(AL353800.10:18548..18641),
complement(AL353800.10:14582..14725),
complement(AL353800.10:9216..11716),
complement(AL353800.10:1566..6816))
/genes="RP3-442A17.1"
```

gene

```
/locus_tag="RP3-442A17.1-001"
join(complement(162776..164384),complement(63851..63979),
complement(47147..47308),complement(27868..28048),
complement(24776..24910),complement(15315..15481),
complement(AL353800.10:33419..33643),
complement(AL353800.10:32105..32339),
complement(AL353800.10:23662..23806),
complement(AL353800.10:19677..19821),
complement(AL353800.10:18548..18641),
complement(AL353800.10:14582..14725),
complement(AL353800.10:5018..6816))
/genes="RP3-442A17.1"
```

mRNA

```
/locus_tag="RP3-442A17.1-002"
join(complement(162776..164384),complement(63851..63979),
complement(47147..47308),complement(27868..28048),
complement(24776..24910),complement(15315..15481),
complement(AL353800.10:33419..33643),
complement(AL353800.10:32105..32339),
complement(AL353800.10:23662..23806),
complement(AL353800.10:19677..19821),
complement(AL353800.10:18548..18641),
complement(AL353800.10:14582..14725),
complement(AL353800.10:9216..11716),
complement(AL353800.10:1566..6816))
/genes="RP3-442A17.1"
```

```
/locus_tag="RP3-442A17.1-001"
/notes="match: ESTs: BE795693 BE900025 BG174636
```

```
match: cDNAs: AF219945 AF219946"
```

```
join(complement(162776..164384),complement(63851..63979),
complement(47147..47308),complement(27868..28048),
complement(24776..24910),complement(15315..15481),
complement(AL353800.10:33419..33643),
complement(AL353800.10:32105..32339),
complement(AL353800.10:23662..23806),
complement(AL353800.10:19677..19821),
complement(AL353800.10:18548..18641),
complement(AL353800.10:14582..14725),
complement(AL353800.10:5018..6816))
/genes="RP3-442A17.1"
```

```
/locus_tag="RP3-442A17.1-002"
```

```
/notes="match: cDNAs: AB037818 AF288480"
```

```
join(complement(162776..163027),complement(63851..63979),
complement(47147..47308),complement(27868..28048),
complement(24776..24910),complement(15315..15481),
complement(AL353800.10:33419..33643),
complement(AL353800.10:32105..32339),
complement(AL353800.10:23662..23806),
complement(AL353800.10:19677..19821),
complement(AL353800.10:18548..18641),
complement(AL353800.10:14582..14725),
complement(AL353800.10:5018..6816))
/genes="RP3-442A17.1"
```

CDS

```
complement(AL353800.10:18548..18641),
complement(AL353800.10:14582..14725),
complement(AL353800.10:9216..11716),
complement(AL353800.10:6700..6816))
/genes="RP3-442A17.1"
/locus_tag="RP3-442A17.1-001"
/standard_name="OTTHUMP0000017502"
/notes="match: proteins: Q9HD22 Q9JIL5"
/codon_start=1
/protein_id="CAI16559.1"
/db_xref="GI:55959499"
/db_xref="GOA:Q5T3M3"
/db_xref="InterPro:IPR000007"
/db_xref="InterPro:IPR001496"
/db_xref="InterPro:IPR001680"
/db_xref="UniProt/TREMBL:Q5T3M3"
/translation="MYAAVEHGPVLCSDSNILCLSWKGRVPSKSEKPKVCRRYEERG
WLATNGRGVGVGVTFTSSHCRDRSTPQRINFNLRHGNSEVLVRWNEPKQKATCDA
DGGIFVMIQTEGRWSVELVNDRGAQVSDFTWSHDGTQALLISYRDGFLVGSVSGQRHW
SSEINLESQITCGIWTDPDQVLFGTADQVIMDCHGRMLAHVLLHESDGLGMSWN
YPIFLVEDSSESDTDDYAPQDGAAPYIPVQNIKPLLTFTSGDLSLMNYYDDL
STTVIRSLGKEVVAQCTQGDLLAVAGMERQTQGLPNGLPKSAMVKNVVRGEHI
FTLDTLVQRPIISICWHRDRLMAGSPALIVRVRHVSLSQLLCCQAIASTLRD
KOVSKITLPPRLCSYLSTAFIPTIKPIPPNNMDFVSPYSAGNELHCTMKRTEDD
PEVGGPCYTLVLELGLVPIILKRRISIKLRFEFVMDPTDTSKPDEIYGNLSITVI
DSCNCSDDSIELSDDDAAKGPKISRASKPKLPRISEARSKPKLPRAAQBLRSVP
RLPLRSPVSGPSLTRRPFEDIQHNYLAQVTSNIWTKFKI VGLAAFLPTNLGAV
IYKTSLLHLQPRQMTIYLPVNRKISMDYINLVPNNVSEDEDDLPVTGASGVENS
PCTVNIPIAPIHSSAGVMSQSIGLQNSLANQVQLDVLNTQTAVGTABHAGDS
ATQYVSNRYSNPGQVIFGSEVMGRIIQNPPLSLPFPQGMQLSTVGGDHDHHL
QKSALALRPTQLAAEGDAVFSAPQEVQVTKINPPPPYGTIPAAFTTAAPPLPP
POPVDVCLTKGDFSLYPTSVHVTPLGYERTITPDSGNVEEVCPRTRMLCSQNTY
TLPGPSSATLRLTATEKKVPQCSATNRLTVPRYSIPTGDPVPYPTASOLAQR
GAAQSDSLIHATLRNNREATLQMAQLADSPRALQLPLAKSKSGGPGVVTOLQAPR
PPALYTCSCSGTGPSSGASLHATASAPASQSSYLSLLSPDSARDTIDVNSAF
PTELALSHQLEKPLRHPPLPEAAVTLKRPYPYQMDLGEDVWPQERTTAQTSGN
PDKLSLMSLQQHLDLPLFTISPKSPASPTATFQTQGMGVYFPGSYNNPPLGVQ
APCSACALSPQFOAQEPAVVLQPLVPSLSYCTLPMPYSGSTCSLQLPVVLPALH
WSYSPACPMQNPQGTLPKPHLVKELVSPADLQSHLGTVMVETADNFOEVLIS
LTSPVPQRTKFKKRLKLSRAEGSVQALTEGKKEARTLSDFNLSLSSPHLG
RRKKVKYSQDKLKSINKNLNTEFQSESEPELFGDELMMNQSGSRKWSKRSRP
RAAGLEAKCRRASEKEDGRGQGVYVMANKQPLMNEATQVYQLDFGGRVTVBSA
KNQFLEGRQVWFGRIDGSAVILDFYPPFSAQAFVALANVTQLK"
```

CDS

```
join(complement(162776..163027),complement(63851..63979),
complement(47147..47308),complement(27868..28048),
complement(24776..24910),complement(15315..15481),
complement(AL353800.10:33419..33643),
complement(AL353800.10:32105..32339),
complement(AL353800.10:23662..23806),
complement(AL353800.10:19677..19821),
complement(AL353800.10:18548..18641),
complement(AL353800.10:14582..14725),
complement(AL353800.10:6794..6816))
/genes="RP3-442A17.1"
/locus_tag="RP3-442A17.1-002"
/standard_name="OTTHUMP0000017503"
/notes="match: proteins: Q9P2F0"
/codon_start=1
/protein_id="CAI16560.1"
/db_xref="GI:55959500"
/translation="MYAAVEHGPVLCSDSNILCLSWKGRVPSKSEKPKVCRRYEERG
WLATNGRGVGVGVTFTSSHCRDRSTPQRINFNLRHGNSEVLVRWNEPKQKATCDA
DGGIFVMIQTEGRWSVELVNDRGAQVSDFTWSHDGTQALLISYRDGFLVGSVSGQRHW
SSEINLESQITCGIWTDPDQVLFGTADQVIMDCHGRMLAHVLLHESDGLGMSWN
YPIFLVEDSSESDTDDYAPQDGAAPYIPVQNIKPLLTFTSGDLSLMNYYDDL
STTVIRSLGKEVVAQCTQGDLLAVAGMERQTQGLPNGLPKSAMVKNVVRGEHI
FTLDTLVQRPIISICWHRDRLMAGSPALIVRVRHVSLSQLLCCQAIASTLRD
KOVSKITLPPRLCSYLSTAFIPTIKPIPPNNMDFVSPYSAGNELHCTMKRTEDD
PEVGGPCYTLVLELGLVPIILKRRISIKLRFEFVMDPTDTSKPDEIYGNLSITVI
DSCNCSDDSIELSDDDAAKGPKISRASKPKLPRISEARSKPKLPRAAQBLRSVP
RLPLRSPVSGPSLTRRPFEDIQHNYLAQVTSNIWTKFKI VGLAAFLPTNLGAV
IYKTSLLHLQPRQMTIYLPVNRKISMDYINLVPNNVSEDEDDLPDGAVWTD"
```

49349..49675

misc_feature

/note="Sequence from overlapping clone AL1161721. Assembly confirmed by restriction digest data."

75390.75577

/note="Sequence from overlapping clone AL353800. Assembly confirmed by restriction digest data."

Join(78483.78534,79069.79159)

/locus tag="RP11-732M18.2-001"

Join(78483.78534,79069.79159)

/locus tag="RP11-732M18.2-001"

/note="match: ESTs: B1463512"

92518

misc_feature

Query Match 80.0%; Score 20; DB 8; Length 180635;

Best Local Similarity 100.0%; Pred. No. 1e+02;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 4 AAATGGGAATCTCAGGCTGA 23

Db 163421 AAATGGGAATCTCAGGCTGA 163402

AC108611 295920 bp DNA linear HTG 08-OCT-2002

Rattus norvegicus clone CH230-292P23, *** SEQUENCING IN PROGRESS

*** 7 unordered pieces.

AC108611

AC108611.4 GI:23195549

HTG; HTGS PHASE1; HTGS DRAFT; HTGS_ENRICHED.

Rattus norvegicus (Norway rat)

ORGANISM

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridea; Muridae; Murinae; Rattus.

1 (bases 1 to 295920)

Muzny,D.Marie., Metzker,M.Lee., Abramson,S., Amin,A., Anguiano,D., Allen,C., Allen,H., Alsbrooks,S., Amin,A., Ayodeji,M., Baca,E., Baden,H., Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H., Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F., Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M., Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E., Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J., Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L., Davila,M.L., Davis,C., Davy-carroll,L., De Anda,C., Dederich,D., Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K., Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G., Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P., Frazer,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M., Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerrero,W., Guevara,W., Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K., Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J., Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M., Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A., Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A., Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,D., Kowar,C., Kowis,C., Kratt,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J., Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J., Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J., Maheshwari,M., Mahindarne,M., Mahmoud,M., Malloy,K., Mangum,A., Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E., Mawhney,S., McLeod,M.P., McNeill,T.Z., Meenen,E., Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S., Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L., Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Nwaokemele,O., Okumuon G., Olarnpunsagoon,A., Pal,S., Parks,K., Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C., Plopper,F., Poindexter,A., Popovic,D., Primus,E., Fu,L.-L., Puafo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F., Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J., Sanders,W., Savary,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,

TITLE JOURNAL

REFERENCE

AUTHORS

TITLE JOURNAL

REFERENCE

AUTHORS

TITLE JOURNAL

COMMENT

Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajls,D., Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J., Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabot,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K., Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J., Warren,K., Wei,X., White,F., Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,X., Zhou,S., Zhao,S., Dunn,D., von Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O., Weinstock,G. and Gibbs,R.A.

Direct Submission

Unpublished

2 (bases 1 to 295920)

Worley,K.C.

Direct Submission

Submitted (31-JAN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 295920)

Rat Genome Sequencing Consortium.

Direct Submission

Submitted (08-OCT-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On Sep 19, 2002 this sequence version replaced gi:21737687.

The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: http://www.hgsc.bcm.tmc.edu/

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GPON

Center clone name: CH230-292P23

----- Summary Statistics

Assembly program: Phrap; version 0.990329

Consensus quality: 252123 bases at least Q40

Consensus quality: 256495 bases at least Q30

Consensus quality: 259425 bases at least Q20

Estimated insert size: 283308; sum-of-contigs estimation

Quality coverage: 3x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)

* NOTE: This sequence may represent more than one clone.

* NOTE: This is a 'working draft' sequence. It currently consists of 7 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

* 1

* 52117: contig of 52117 bp in length

* 52218: gap of unknown length

* 270693: contig of 218476 bp in length

* 270694: gap of unknown length

* 270793: gap of unknown length

* 272443: contig of 1350 bp in length

* 272144: gap of unknown length

* 272564: contig of 2321 bp in length

* 274565: gap of unknown length


```

* 274665 276601: contig of 1937 bp in length
* 276502 276701: gap of unknown length
* 276702 288825: contig of 12124 bp in length
* 288826 288925: gap of unknown length
* 288926 295920: contig of 6995 bp in length.
FEATURES
    source
        1. 295920
            /organism="Rattus norvegicus"
            /mol_type="genomic DNA"
            /db_xref="taxon:10116"
            /clone="CH230-292P23"
            /notes="wgs contig"
        1. 1408
            /estimated_length=unknown
        52118..52117
            /estimated_length=unknown
        270694..270793
            /estimated_length=unknown
        272144..272243
            /estimated_length=unknown
        274565..274664
            /estimated_length=unknown
        276602..276701
            /estimated_length=unknown
        288826..288925
            /estimated_length=unknown
ORIGIN
Query Match      80.0%; Score 20; DB 14; Length 295920;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAAAACTGAATCTCAGGC 20
   |||||
Db 120957 AAAAACTGAATCTCAGGC 120976

RESULT 25
AC025192      134514 bp DNA linear HTG 27-MAR-2003
LOCUS Homo sapiens chromosome 8 clone RP11-509E2 map 8, 3 unordered
DEFINITION pieces.
ACCESSION AC025192
VERSION AC025192.4 GI:18875263
KEYWORDS HTG; HTGS PHASE1; HTGS_FULLTOP; HTGS_CANCELLED.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
1 (bases 1 to 134514)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens chromosome 8, clone RP11-509E2
Unpublished
2 (bases 1 to 134514)
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
Boguslavskiy,L., Boukhgalter,B., Brown,A., Burkett,G.,
Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
Collamore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., LaRoque,K., Lamazares,R., Landers,T., Lehoczy,J.,
Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,

```

```

Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (07-MAR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 134514)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L.,
Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collamore,A.,
Cooke,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gordin,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., LaRoque,K., Lamazares,R.,
Landers,T., Lehoczy,J., Levine,R., Lindblad-Toh,K., Liu,G.,
Maclean,C., Macdonald,P., Major,J., Marquis,N., Meneus,L.,
McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L.,
Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (27-MAR-2003) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Feb 25, 2002 this sequence version replaced gi:123113839.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence.submissions@genome.wi.mit.edu
----- Project Information
Center project name: L5592
Center clone name: 509_E_2
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 17787: contig of 17787 bp in length
* 17788 17887: gap of 100 bp
* 17888 97189: contig of 79301 bp in length
* 97189 97288: gap of 100 bp
* 97289 134514: contig of 37226 bp in length.
FEATURES
    Location/Qualifiers
        1. 134514
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            /chromosome="8"
            /map="8"
            /clone="rp11-509E2"
            /clone_lib="RP11-11 Human Male BAC"
            17788..17887
                /estimated_length=100
            97189..97288
                /estimated_length=100
ORIGIN
Query Match      79.2%; Score 19.8; DB 14; Length 134514;

```

```
Best Local Similarity 91.3%; Pred. No. 1.3e+02;
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AAAAAGTGAATCTCAGGCTGAG 24
    ||||| ||||| ||||| |||||
Db 128346 AAAGACTGGAACCTCAGGCTGAG 128368

RESULT 26
AC069005      162560 bp      DNA      linear      HTG 16-JUL-2000
LOCUS      Homo sapiens chromosome 8 clone RP11-712115, WORKING DRAFT
DEFINITION      AC069005
SEQUENCE, 33 unordered pieces.
ACCESSION      AC069005
VERSION        AC069005.3 GI:8844180
KEYWORDS      HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE        Homo sapiens (human)
ORGANISM      Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
              Homiidae; Homo.
REFERENCE      1 (bases 1 to 162560)
AUTHORS      Waterston,R.H.
TITLE        The sequence of Homo sapiens clone
JOURNAL      Unpublished
REFERENCE      2 (bases 1 to 162560)
AUTHORS      Waterston,R.H.
TITLE        Direct Submission
JOURNAL      Submitted (16-MAY-2000) Genome Sequencing Center, Washington
              University School of Medicine, 444 Forest Park Parkway, St. Louis,
              MO 63108, USA
COMMENT      On Jun 30, 2000 this sequence version replaced gi:8469066.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H NH0712115
----- Summary Statistics -----
Sequencing vector: plasmid; 100%
Chemistry: Dye-primer ET; 100% of reads
Assembly: Phrap; version 0.990319
Consensus quality: 145600 bases at least Q40
Consensus quality: 151065 bases at least Q30
Consensus quality: 153390 bases at least Q20
Insert size: 173000; agarose-fp
Insert size: 159360; sum-of-contigs
Quality coverage: 3.35 in Q20 bases; agarose-fp
Quality coverage: 3.73 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 33 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 1483: contig of 1483 bp in length
* 1484 1583: gap of unknown length
* 1584 3194: contig of 1611 bp in length
* 3195 3294: gap of unknown length
* 3295 4472: contig of 1178 bp in length
* 4473 4572: gap of unknown length
* 4573 6120: contig of 1548 bp in length
* 6121 6220: gap of unknown length
* 6221 8046: contig of 1826 bp in length
* 8047 8146: gap of unknown length
* 8147 9901: contig of 1755 bp in length
* 9902 10001: gap of unknown length

FEATURES
source
1..162560
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="8"
/clone="RP11-712115"
1484..1583
/estimated_length=unknown
3195..3294
/estimated_length=unknown
4473..4572
/estimated_length=unknown
6121..6220
/estimated_length=unknown
8047..8146
/estimated_length=unknown
9902..10001
/estimated_length=unknown
12501..12600
/estimated_length=unknown

10002 12500: contig of 2499 bp in length
12501 12600: gap of unknown length
12601 15266: contig of 2666 bp in length
15267 15366: gap of unknown length
15367 17549: contig of 2183 bp in length
17550 17650: gap of unknown length
17651 20378: contig of 2729 bp in length
20379 20478: gap of unknown length
20479 23345: contig of 2867 bp in length
23346 27122: contig of 3677 bp in length
27123 30048: contig of 2826 bp in length
30049 30148: gap of unknown length
30149 32220: contig of 2072 bp in length
32221 35967: contig of 3647 bp in length
35968 36068: gap of unknown length
36069 39730: contig of 3663 bp in length
39731 43447: contig of 3617 bp in length
43448 43548: gap of unknown length
43549 47467: contig of 3920 bp in length
47468 47568: gap of unknown length
47569 51881: contig of 4313 bp in length
51882 51980: gap of unknown length
51981 57315: contig of 5335 bp in length
57316 61833: contig of 4418 bp in length
61834 61934: gap of unknown length
61935 65951: contig of 4018 bp in length
65952 66051: gap of unknown length
66052 70869: contig of 4818 bp in length
70870 70969: gap of unknown length
70970 77012: contig of 6043 bp in length
77013 82785: contig of 5673 bp in length
82786 82885: gap of unknown length
82886 88405: contig of 5520 bp in length
88406 88506: gap of unknown length
88507 93768: contig of 5262 bp in length
93769 93868: contig of 5705 bp in length
93869 99572: gap of unknown length
99573 109490: contig of 9818 bp in length
109491 109590: gap of unknown length
109591 119557: contig of 9967 bp in length
119558 119657: gap of unknown length
119658 131547: contig of 11890 bp in length
131548 131647: gap of unknown length
131648 148258: contig of 16611 bp in length
148259 148358: gap of unknown length
148359 162560: contig of 14202 bp in length.

Location/Qualifiers
1..162560
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="8"
/clone="RP11-712115"
1484..1583
/estimated_length=unknown
3195..3294
/estimated_length=unknown
4473..4572
/estimated_length=unknown
6121..6220
/estimated_length=unknown
8047..8146
/estimated_length=unknown
9902..10001
/estimated_length=unknown
12501..12600
/estimated_length=unknown
```

```

/estimated_length=unknown
15267..15366
/estimated_length=unknown
17550..17649
/estimated_length=unknown
20379..20478
/estimated_length=unknown
23346..23445
/estimated_length=unknown
27123..27222
/estimated_length=unknown
30049..30148
/estimated_length=unknown
32221..32320
/estimated_length=unknown
35968..36067
/estimated_length=unknown
39731..39830
/estimated_length=unknown
43448..43547
/estimated_length=unknown
47468..47567
/estimated_length=unknown
51881..51980
/estimated_length=unknown
57316..57415
/estimated_length=unknown
61834..61933
/estimated_length=unknown
65952..66051
/estimated_length=unknown
70870..70969
/estimated_length=unknown
77013..77112
/estimated_length=unknown
82786..82885
/estimated_length=unknown
88406..88505
/estimated_length=unknown
93768..93867
/estimated_length=unknown
99573..99672
/estimated_length=unknown
109491..109590
/estimated_length=unknown
119558..119657
/estimated_length=unknown
131548..131647
/estimated_length=unknown
148259..148358
/estimated_length=unknown

```

ORIGIN

```

Query Match 79.2%; Score 19.8; DB 14; Length 162560;
Best Local Similarity 91.3%; Pred. No. 1.3e+02;
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AAAAATCGGAATCTCAGGCTGAG 24
   |||||
Db 84983 AGAATCGGAATCTCAGGCTGAG 85005
   |||||

RESULT 27
AC103719
LOCUS Homo sapiens chromosome 8, clone RP11-421P23, complete sequence.
DEFINITION AC103719
ACCESSION AC103719
VERSION HTG.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

```

REFERENCE
AUTHORS

1 (bases 1 to 167878)
Birren,B., Nusbaum,C. and Lander,E.
Homo sapiens chromosome 8, clone RP11-421P23

JOURNAL
AUTHORS

Unpublished
2 (bases 1 to 167878)

Hominidae; Homo.

Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Boguslavskiy,L., Boukhgalter,B.,
Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Hulme,W., Iliev,I., Johnson,R.,
Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K.,
Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G.,
MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J.,
Meneus,L., Mihova,T., Mienga,V., Murphy,T., Naylor,J., Nguyen,C.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission

Submitted (29-NOV-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA

3 (bases 1 to 167878)

TITLE
JOURNAL

Unpublished

REFERENCE
AUTHORS

Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B.,
Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A., S.,
Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J.,
Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hafez,N.,
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
Lindblad-Toh,K., Liu,G., MacLean,C., Macdonald,P., Major,J.,
Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T.,
Mienga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J.,
Peterson,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R.,
Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Talamas,J., Testaye,S., Theodore,J., Topham,K.,
Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X.,
Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission

Submitted (03-JAN-2003) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA

4 (bases 1 to 167878)

TITLE
JOURNAL

Unpublished

REFERENCE
AUTHORS

Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B.,
Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J.,
Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hafez,N.,
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
Lindblad-Toh,K., Liu,G., MacLean,C., Macdonald,P., Major,J.,
Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T.,
Mienga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J.,
Peterson,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R.,
Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Talamas,J., Testaye,S., Theodore,J., Topham,K.,
Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X.,
Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission

Baldwin, J., Barna, N., Beckerly, R., Boguslavskiy, L., Boukhgalter, B., Brown, A., Castle, A., Collangelo, M., Collins, S., Collamore, A., Cooke, P., DeArellano, K., Dewar, K., Domino, M., Donelan, L., Doyle, M., Ferreira, P., FitzHugh, W., Forrest, C., Funke, R., Gage, D., Galagan, J., Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L., Howland, J. C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J., Lehoczy, J., Lieu, C., Locke, K., Macdonald, P., Marquis, N., McEwan, P., McGurk, A., McKernan, K., McLaughlin, J., Meldrim, J., Morrow, J., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P., Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Tesfaye, S., Tirrell, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X., Wyman, D., Ye, W. J., Zimmer, A. and Zody, M.

TITLE

Submitted (16-NOV-1999) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE

3 (bases 1 to 189662)

AUTHORS

Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, S., Barna, N., Bastien, V., Boguslavskiy, L., Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Choepel, Y., Colangelo, M., Collins, S., Collamore, A., Cooke, P., DeArellano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Karatas, A., LaRocque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G., MacLean, C., Macdonald, P., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., McPheters, R., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schuback, R., Seaman, S., Severy, P., Sougnez, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Straus, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE

Submitted (01-MAY-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

COMMENT

On May 1, 2001 this sequence version replaced gi:12313808.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information
Center project name: L2466
Center clone name: 369_E_15

FEATURES

source
Location/Qualifiers
1. .189662
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="8"
/map="8"
/clone="RP11-369E15"
/clone_lib="RPC1-11 Human Male BAC"
1. .100
/rpt_family="MLT1A1"
1594. .1741
/rpt_family="GA-rich"
complement(2782. .2957)
/rpt_family="MIR"
3311. .3471
/rpt_family="MIR"
3964. .4067
/rpt_family="MIR"

repeat_region
complement(4528. .4771)
/rpt_family="L2"
4776. .4871
/rpt_family="MER4A"
4869. .5223
/rpt_family="MER4A"
complement(5255. .5765)
/rpt_family="L2"
complement(5805. .6003)
/rpt_family="THE1A"
6004. .6035
/rpt_family="TTA)n"
complement(6151. .6319)
/rpt_family="SVA"
complement(6320. .6468)
/rpt_family="THE1A"
6722. .8751
/rpt_family="TTTAA)n"
9155. .9553
/rpt_family="MSTB"
9599. .9982
/rpt_family="THE1C"
complement(10113. .10320)
/rpt_family="AluY"
10822. .10970
/rpt_family="MIR3"
11210. .11215
/note="<30 qual SINGL region"
11794. .12080
/rpt_family="AluSx"
12974. .13010
/rpt_family="(TG)n"
complement(13091. .13304)
/rpt_family="MER106A"
13481. .13858
/rpt_family="MLT1C"
14025. .14260
/rpt_family="L2"
14584. .14679
/rpt_family="MIR"
14721. .14825
/rpt_family="L2"
15167. .15332
/rpt_family="MER33"
15436. .15788
/rpt_family="L2"
complement(16078. .16165)
/rpt_family="AluJ/FRAM"
16532. .17094
/rpt_family="MLT1F1"
17653. .17674
/rpt_family="(TAA)n"
complement(17675. .17892)
/rpt_family="MER7C"
complement(17897. .18650)
/rpt_family="LTR17"
complement(18651. .18996)
/rpt_family="MER7C"
18999. .19103
/rpt_family="(TAA)n"
complement(19203. .19283)
/rpt_family="MIR"
20464. .20523
/rpt_family="A-rich"
complement(21279. .21902)
/rpt_family="MER41B"
21922. .22092
/rpt_family="MIR"
22249. .22373
/rpt_family="MIR"
23956. .24058
/rpt_family="L2"
complement(24069. .24198)

```

repeat_region /rpt_family="MIR"
25149 .25213
/rpt_family="L2"
repeat_region /rpt_family="L2"
complement(25783. .26130)
repeat_region /rpt_family="MUT11"
complement(26174. .26539)
repeat_region /rpt_family="MUT11"
complement(28187. .28528)
repeat_region /rpt_family="AluY"
complement(28792. .29263)
repeat_region /rpt_family="Charlie8"
complement(30652. .30730)
repeat_region /rpt_family="MIR53"
complement(30986. .31117)
repeat_region /rpt_family="LIME1"
31113. .31556
/rpt_family="LIME2"
31590. .32174
repeat_region /rpt_family="LIME2"
32162. .32588
/rpt_family="LIME3"
33242. .33490
repeat_region /rpt_family="MIR"
complement(34416. .34803)
repeat_region /rpt_family="MSTB"
complement(35044. .35350)
repeat_region /rpt_family="AluSg"
36407. .36695
/rpt_family="AluSx"
complement(37670. .37790)
repeat_region /rpt_family="FLAM A"
complement(38065. .38132)
repeat_region /rpt_family="L2"
complement(38504. .38590)
repeat_region /rpt_family="MIR"
39044. .39228
/rpt_family="MIR3"
complement(40317. .40473)

Query Match 79.2%; Score 19.8; DB 8; Length 189662;
Best Local Similarity 91.3%; Pred. No. 1.3e+02;
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AAAAATCGAATCTCAGGCTGAG 24
Db 165165 AGAATCTGGAATCTCAGGCTGAG 165187
|||||
|||||

RESULT 29
AC087221/c 203690 bp DNA linear HTG 24-MAY-2002
LOCUS Homo sapiens chromosome 8 clone RP11-712115 map 8, WORKING DRAFT
DEFINITION SEQUENCE, 34 ordered pieces.
ACCESSION AC087221
VERSION AC087221.2 GI:211166223
KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE 1 (bases 1 to 203690)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Allen,N., Anderson,S.,
TITLE Camarata,J., Campopiano,A., Choepel,Y., Colangelo,M., Collins,S.,
JOURNAL Collimore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,
REFERENCE Dodg,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J.,
2 (bases 1 to 203690)
AUTHORS Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Headford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,

```

```

Jones, C., Karatas,A., LaRocque,K., Lamazares,R., Landers,T.,
Lehoczy,J., Levine,R., Liu,G., Maclean,C., Macdonald, P.,
Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K.,
McNeeters,R., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H.,
O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R.,
Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M.,
Roy,A., Santos,R., Schauer,S., Schuback,R., Seaman,S., Severy,P.,
Sougne,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A.,
Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J.,
Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (18-DEC-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 203690)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L.,
Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collimore,A.,
Cooke,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., FitzGerald,M., FitzHugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Gird,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I.,
Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K.,
Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Lindblad-Toh,K.,
Liu,G., Maclean,C., Macdonald,P., Major,J., Marquis,N.,
Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Meldrim,J.,
Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
Nicol,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C.,
Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S.,
Schuback,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S.,
Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (24-MAY-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On May 24, 2002 this sequence version replaced gi:11875303.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submission@genome.wi.mit.edu
----- Project Information
Center project name: L11638
Center clone name: 712 I 15
----- Summary Statistics
Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 190084 bases at least Q40
Consensus quality: 196705 bases at least Q30
Consensus quality: 199369 bases at least Q20
Insert size: 176000; agarose-ep
Insert size: 200390; sum-of-contigs
Quality coverage: 9.6 in Q20 bases; agarose-ep
Quality coverage: 8.5 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 34 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submittor.

```

```

TITLE
JOURNAL
COMMENT

```

/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="8"
/map="8"
/clone_lib="RPCI-II Human Male BAC"
1..460
/note="assembly_fragment"
clone_end:SP6
vector_side:left"
461..560
/estimated_length=100
561..1196
/note="assembly_fragment"
1197..1296
/estimated_length=100
1297..2122
/note="assembly_fragment"
2123..2222
/estimated_length=100
2223..2869
/note="assembly_fragment"
2870..2969
/estimated_length=100
2970..3620
/note="assembly_fragment"
3621..3720
/estimated_length=100
3721..4314
/note="assembly_fragment"
4315..4414

misc_feature
1..460

gap
461..560

misc_feature
561..1196

gap
1197..1296

misc_feature
1297..2122

gap
2123..2222

misc_feature
2223..2869

gap
2870..2969

misc_feature
2970..3620

gap
3621..3720

misc_feature
3721..4314

gap
4315..4414

Query Match 79.2%; Score 19.8; DB 14; Length 203690;
Best Local Similarity 91.3%; Pred.No.1.3e+02;
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AAAAACTGGAATCTCAGGCTGAG 24
DB 153635 AGAATCTGGAATCTCAGGCTGAG 153613

AC156131 225793 bp DNA linear HTG 01-JUL-2005
Bos taurus clone CH240-47A8, *** SEQUENCING IN PROGRESS ***, 18
unordered pieces.
AC156131
AC156131.2 GI:68265537
HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
Bos taurus (cow)
Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
Pecora; Bovidae; Bovinae; Bos.
1 (bases 1 to 225793)
Muzny, D. Marie., Metzker, M. Lee., Abramson, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Alshrocks, S., Amin, A., Anguiano, D.,
Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,
Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
Egan, A., Escoto, M., Eugene, C., Evans, C. A., Fallis, T., Fan, G.,
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
Gebrgeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K.,
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,

* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
1
460: contig of 460 bp in length
561: gap of 100 bp
1196: contig of 636 bp in length
1296: gap of 100 bp
2122: contig of 826 bp in length
2222: gap of 100 bp
2869: contig of 647 bp in length
3620: contig of 651 bp in length
3720: gap of 100 bp
4314: contig of 594 bp in length
4414: gap of 100 bp
5152: contig of 738 bp in length
5252: gap of 100 bp
6209: contig of 957 bp in length
6309: gap of 100 bp
7103: contig of 794 bp in length
7203: gap of 100 bp
7937: contig of 734 bp in length
8037: gap of 100 bp
9106: contig of 969 bp in length
9906: contig of 800 bp in length
10006: gap of 100 bp
11070: contig of 1064 bp in length
11170: gap of 100 bp
12297: contig of 1127 bp in length
12397: gap of 100 bp
13415: contig of 1018 bp in length
13515: gap of 100 bp
14231: contig of 716 bp in length
14331: gap of 100 bp
15596: contig of 1265 bp in length
15696: gap of 100 bp
16462: contig of 766 bp in length
16562: gap of 100 bp
17742: contig of 1180 bp in length
17842: gap of 100 bp
18857: contig of 1015 bp in length
18957: gap of 100 bp
20490: contig of 1533 bp in length
20590: gap of 100 bp
22210: contig of 1620 bp in length
22310: gap of 100 bp
23851: contig of 1541 bp in length
23951: gap of 100 bp
25684: contig of 1733 bp in length
25784: gap of 100 bp
27006: contig of 1222 bp in length
27106: gap of 100 bp
28616: contig of 1510 bp in length
28716: gap of 100 bp
30184: contig of 1468 bp in length
30284: gap of 100 bp
31303: contig of 1019 bp in length
31403: gap of 100 bp
32917: contig of 1514 bp in length
33017: gap of 100 bp
42030: contig of 9013 bp in length
42130: gap of 100 bp
49012: contig of 6882 bp in length
49112: gap of 100 bp
62240: contig of 13128 bp in length
62340: gap of 100 bp
86795: contig of 24455 bp in length
86796: gap of 100 bp
86896 203690: contig of 116795 bp in length.
Location/Qualifiers
1..203690
/organism="Homo sapiens"

Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M., Hollins,B., Howells,S., Huiyk,S., Hume,J., Idlebird,D., Jackson,A., Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A., Karpathy,S., Kelly,S., Khan,Z., King,L., Kovar,C., Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J., Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J., Loresushewa,L., Loulsegred,H., Lozado,R.J., Lu,X., Ma,J., Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A., Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E., Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E., Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S., Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L., Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K., Pasternak,S., Paul,H., Perez,A., Perez,L., Pfankoch,C., Popper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L., Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F., Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J., Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H., Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D., Sneed,A., Sodergren,B., Song,X.-Z., Sorelle,R., Sober,V., Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Umani,K., Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J., Wang,O., Wang,S., Warren,J., Warren,R., Wei,X., White,F., Williams,G., Willson,R., Wlciyck,R., Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O., Weinstein,G. and Gibbs,R.A.

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished
2 (bases 1 to 225793)

Worley,K.C.

Direct Submission

Submitted (25-JAN-2005) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 225793)

Direct Submission

Submitted (01-JUN-2005) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On Jun 28, 2005 this sequence version replaced gi:58082168.

The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: FCQG

Center clone name: CH240-47A8

----- Summary Statistics

Assembly program: Atlas 3.0;

Consensus quality: 218147 bases at least Q40

Consensus quality: 219918 bases at least Q30

Consensus quality: 221497 bases at least Q20

Estimated insert size: 222678; sum-of-contigs estimation

Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently consists of 18 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1
6241: contig of 6241 bp in length
6291: gap of 50 bp
6292: 19465: contig of 13174 bp in length
19466: 19570: gap of 105 bp
19571: 55981: contig of 36411 bp in length
55982: 56031: gap of 50 bp
56032: 73696: contig of 17665 bp in length
73697: 73746: gap of 50 bp
73747: 74761: contig of 1015 bp in length
74762: 75217: gap of 456 bp
75218: 77049: contig of 1832 bp in length
77050: 77149: gap of unknown length
101400: contig of 24251 bp in length
101401: 101874: gap of 474 bp
101875: 129165: contig of 27291 bp in length
129166: 129265: gap of unknown length
129266: 161426: contig of 32161 bp in length
161427: 16476: gap of 50 bp
16477: 179575: contig of 18099 bp in length
179576: 179625: gap of 50 bp
179626: 202028: contig of 22403 bp in length
202029: 202078: gap of 50 bp
202079: 204586: contig of 2508 bp in length
204587: 204687: gap of unknown length
204688: 214263: contig of 9577 bp in length
214264: 214441: gap of 178 bp
214442: 219209: contig of 4768 bp in length
219210: 219309: gap of unknown length
219310: 220717: contig of 1408 bp in length
220718: 220817: gap of unknown length
220818: 222538: contig of 1721 bp in length
222539: 222638: gap of unknown length
222639: 224291: contig of 1653 bp in length
224292: 224391: gap of unknown length
224392: 225793: contig of 1402 bp in length.

FEATURES

source

1..225793
/organism="Bos taurus"
/mol_type="genomic DNA"
/db_xref="taxon:9913"
/clone="CH240-47A8"
6242..6291
/estimated_length=50
19466..19570
/estimated_length=105
55982..56031
/estimated_length=50
73697..73746
/estimated_length=50
74762..75217
/estimated_length=456
77050..77149
/estimated_length=unknown
101401..101874
/estimated_length=474
129166..129265
/estimated_length=unknown
161427..161476
/estimated_length=50
179576..179625
/estimated_length=50
202029..202078
/estimated_length=50

gap

gap

gap

gap

gap

gap

gap

gap

gap

gap

gap

/estimated_length=50
 204587 .204686
 /estimated_length=unknown
 214264 .214441
 /estimated_length=178
 219210 .219309
 /estimated_length=unknown
 220718 .220817
 /estimated_length=unknown
 222539 .222638
 /estimated_length=unknown
 224292 .224391
 /estimated_length=unknown

ORIGIN

Query Match 79.2%; Score 19.8; DB 14; Length 225793;
 Best Local Similarity 91.3%; Pred. No. 1.3e+02;
 Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

 Qy 2 AAAAAGTGGAAATCTCAGGCTGAG 24
 Db 123076 ACAAAGTGGAAATCTCAGGCTGAG 123054

RESULT 31

AC152300/c
 LOCUS AC152300 227115 bp DNA linear HTG 01-JUL-2005
 DEFINITION Bos taurus clone CH240-1K11, *** SEQUENCING IN PROGRESS ***, 29
 unorderd pieces.
 AC152300
 VERSION AC152300.3 GI:68227212
 KEYWORDS HTG; HTGS PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
 SOURCE Bos taurus (Cow)

REFERENCE

AUTHORS
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
 Pecora; Bovidae; Bovinae; Bos.
 1 (bases 1 to 227115)
 Muzny, D., Maric, M., Metzker, M., Lee, S., Abramson, S., Adams, C., Alder, J.,
 Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
 Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
 Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
 Blawie, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
 Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
 Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Chen, A., Chu, J.,
 Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
 Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
 Davila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
 Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
 Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
 Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G.,
 Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
 Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
 Gebregorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
 Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K.,
 Harvey, Y., Havlak, P., Haves, A., Henderson, N., Hernandez, J.,
 Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hognes, M.,
 Hollins, B., Howells, S., Huly, S., Hume, J., Idlebird, D., Jackson, A.,
 Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
 Karpachy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
 Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
 Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
 Loresu, H., Loussegod, H., Lozada, R.J., Lu, X., Ma, J.,
 Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A.,
 Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
 Manthey, S., McLeod, M.P., McNeill, T.Z., Meenen, E.,
 Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
 Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,
 Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
 Nwokeneme, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K.,
 Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,
 Plopper, P., Poinexter, A., Popovic, D., Primus, E., Pu, L.-L.,
 Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R.,

TITLE JOURNAL
 REFERENCE
 AUTHORS
 TITLE JOURNAL
 REFERENCE
 AUTHORS
 TITLE JOURNAL

COMMENT

On Jun 26, 2005 this sequence version replaced gi:58038108.
 The sequence in this assembly is a combination of BAC based reads
 and whole genome shotgun sequencing reads assembled using Atlas
 (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
 in the feature table below represents a scaffold in the Atlas
 assembly (a 'contig-scaffold'). Within each contig-scaffold,
 individual sequence contigs are ordered and oriented, and separated
 by sized gaps filled with Ns to the estimated size. The sequence
 may extend beyond the ends of the clone and there may be sequence
 contigs within a contig-scaffold that consist entirely of whole
 genome shotgun sequence reads. Both end sequences and whole genome
 shotgun sequence only contigs will be indicated in the feature
 table.

----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: http://www.hgsc.bcm.tmc.edu/
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: FAXC
 Center clone name: CH240-1K11
 ----- Summary Statistics
 Assembly program: Atlas 3.0;
 Consensus quality: 213949 bases at least Q40
 Consensus quality: 216851 bases at least Q30
 Consensus quality: 219286 bases at least Q20
 Estimated insert size: 214671; sum-of-contigs estimation
 Quality coverage: 5x in Q20 bases; sum-of-contigs estimation

 * NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 29 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 * 1 3732: contig of 3732 bp in length
 * 3733 3782: gap of 50 bp
 * 3783 7703: contig of 3921 bp in length
 * 7704 7804: gap of unknown length
 * 7804 10239: contig of 2436 bp in length
 * 10240 10289: gap of 50 bp

 This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one pUC18 subclone; and the assembly was confirmed by restriction digest.

Neighboring sequence information:
 This clone is overlapped by GSI-25117, GSI-287C17.

 Sequence Quality Assessment:

This entry has been annotated with sequence quality estimates computed by the Phrap assembly program. All manually edited bases have been reduced to quality zero. Quality levels above 40 are expected to have less than 1 error in 10,000 bp.

Base-by-base quality values are not generally visible from the GenBank flat file format but are available as part of this entry's ASN.1 file.

FEATURES

source

Location/Qualifiers

1. .234429
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="8"

source

/map="p21-p11"
 /clone="SCB-212E3"
 218471. .234429
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"

source

/note="overlapping clone"
 233972. .234429
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /clone="GSI-287C17"
 /note="overlapping clone"

misc_feature

1. .1527
 /note="single stranded/single chemistry region"

misc_feature

1752. .1892
 /note="single stranded/single chemistry region , pcr product sequence only"

misc_feature

1999. .2209
 /note="single stranded/single chemistry region"

misc_feature

2405. .2904
 /note="single stranded/single chemistry region"

misc_feature

4045. .4520
 /note="single stranded/single chemistry region"

misc_feature

4909. .5465
 /note="single stranded/single chemistry region"

misc_feature

5006
 /note="low quality region"

misc_feature

5074. .5076
 /note="low quality region"

misc_feature

5958. .6158
 /note="single stranded/single chemistry region"

misc_feature

8506. .9433
 /note="single stranded/single chemistry region"

misc_feature

12839. .12949
 /note="single stranded/single chemistry region"

misc_feature

19289. .19505
 /note="single stranded/single chemistry region"

misc_feature

22256. .22617
 /note="single stranded/single chemistry region"

misc_feature

22915. .23104
 /note="single stranded/single chemistry region"

misc_feature

25717. .25784
 /note="single stranded/single chemistry region"

misc_feature

25979. .26519
 /note="single stranded/single chemistry region"

misc_feature

27780. .27857
 /note="single stranded/single chemistry region"

misc_feature

28946. .28979
 /note="single stranded/single chemistry region"

misc_feature

32258. .32444
 /note="single stranded/single chemistry region"

misc_feature

33909. .34058
 /note="single stranded/single chemistry region"

misc_feature

34482. .34564
 /note="single stranded/single chemistry region"

misc_feature

44437. .44756
 /note="single stranded/single chemistry region"

misc_feature

53510. .53587
 /note="single stranded/single chemistry region"

misc_feature

55018. .55027
 /note="single stranded/single chemistry region"

misc_feature

56892. .57298
 /note="single stranded/single chemistry region"

misc_feature

59734. .59817
 /note="single stranded/single chemistry region"

misc_feature

60277. .60529
 /note="single stranded/single chemistry region"

misc_feature

60712. .60875
 /note="single stranded/single chemistry region"

misc_feature

72531. .72795
 /note="single stranded/single chemistry region"

misc_feature

72810. .72863
 /note="single stranded/single chemistry region"

misc_feature

74251. .74328
 /note="single stranded/single chemistry region"

misc_feature

75828
 /note="low quality region"

misc_feature

76079. .76824
 /note="single stranded/single chemistry region"

misc_feature

81807. .81875
 /note="single stranded/single chemistry region"

misc_feature

83588. .83920
 /note="single stranded/single chemistry region"

misc_feature

86322. .86818
 /note="single stranded/single chemistry region"

misc_feature

88445. .88454
 /note="single stranded/single chemistry region"

misc_feature

89326. .89651
 /note="single stranded/single chemistry region"

misc_feature

90582. .90927
 /note="single stranded/single chemistry region"

misc_feature

95744. .95772
 /note="single stranded/single chemistry region"

misc_feature

98134. .98199
 /note="single stranded/single chemistry region"

misc_feature

98571. .99114
 /note="single stranded/single chemistry region"

misc_feature

102816. .102891
 /note="single stranded/single chemistry region"

misc_feature

102915. .103388
 /note="single stranded/single chemistry region"

misc_feature

103538. .103624
 /note="single stranded/single chemistry region"

misc_feature

103674. .103863
 /note="single stranded/single chemistry region"

misc_feature

108658. .108801
 /note="single stranded/single chemistry region"

misc_feature

110323. .110378
 /note="single stranded/single chemistry region"

misc_feature

114946. .115062
 /note="single stranded/single chemistry region"

misc_feature

115984. .116248
 /note="single stranded/single chemistry region"

misc_feature

118198. .118280
 /note="single stranded/single chemistry region"

misc_feature

119071. .119213
 /note="single stranded/single chemistry region"

Query Match 79.2%; Score 19.8; DB 8; Length 234429;
Best Local Similarity 91.3%; Pred. No. 1.3e+02;
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AAAACTGGAATCTCAGGCTGAG 24
||| ||||| ||||| ||||| |||||
DB 201360 AAAGACTGGAATCTCAGGCTGAG 201338

RESULT 33
AC013717/c
LOCUS 168813 bp DNA linear PRI 15-APR-2005
DEFINITION Homo sapiens BAC clone RP11-2415 from 2, complete sequence.
AC013717
VERSION AC013717.8 GI:14190753
HTG.
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominae; Homo.
REFERENCE
1 (bases 1 to 168813)
Maupin,R. and Edwards,J.
The sequence of Homo sapiens BAC clone RP11-2415
Unpublished (2001)
REFERENCE
2 (bases 1 to 168813)
Waterston,R.H.
Direct Submission
Submitted (13-NOV-1999) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE
3 (bases 1 to 168813)
Waterston,R.H.
Direct Submission
Submitted (07-NOV-2001) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE
5 (bases 1 to 168813)
Wilson,R.K.
Direct Submission
Submitted (15-APR-2005) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
On May 23, 2001 this sequence version replaced gi:13431081.

Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu
Contact: submissions@wustl.edu

Center project name: H_NH0024105

NOTICE:
This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by
restriction digest.
MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. Wes Warren,
Department of Genetics, Washington University, St. Louis MO. For
additional information about the map position of this sequence, see

http://genome.wustl.edu

SOURCE INFORMATION:

The RPCI-11 human BAC library was made from the blood of one male
donor, as described by Oseogawa,K., Woon,P.Y., Zhao,B., Frengen,E.,
Tateno,M., Catanese,J.J. and de Jong,P.J. (1998) An improved
approach for construction of bacterial artificial chromosome
libraries. Genomics 51:1-8. The clone may be obtained either from
Research Genetics, Inc. (http://www.resgen.com) or Pieter de Jong
and coworkers at http://www.chori.org
VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-559M23; the clone sequenced
to the right is RP11-194L1. Actual start of this clone is at base
position 1 of RP11-2415, actual end is at base position 168813 of
RP11-2415.

FEATURES
source

Location/Qualifiers
1..168813
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="2"
/clone="RP11-2415"
/clone_lib="RPCI-11"
8128..38413
/gene="PPM1B"
join(8128..8987,16152..16269,24909..25020,25422..25479,
37355..38413)
/gene="PPM1B"
join(8142..8987,16152..16269,24909..25020,25422..25479,
37355..37660)
/gene="PPM1B"
/note="Homo sapiens protein phosphatase 1B (formerly 2C),
magnesium-dependent, beta isoform (PPM1B), transcript
variant 1, mRNA.; H_NH0024105.1
This gene was based on gi(29558349)
Continued from H_NH0559M23.1"

gene

mRNA

CDS

/product="unknown"
/protein_id="AA88954.1"
/db_xref="GI:62630209"
/translation="MGAFIDPKTEKHNAGNGRLRYGLSSMQGWRVEMEDAHTAVV
GIPGLEWSPFAVDGHSRVANYCSTHLEHTTTFEDFRAAGKSGSALELSVNV
KNGITGFKIDVEMRNFSLRNGMDRSGTAGVMIQPKHYFINCDSRAVLVRNG
QVCFSTQHKPCNPREKRIQAGSGVMIQVNGSLAYSRALGDYDKVDGKGPTEQ
LVSPEPEYILRAEDEFILACDGIWVMSNELCYKSRLEVSDDLNVNCVV
DTCLEKGRDNMSIVLCFNSAPKVSDEAVKSDSELDKHLSEKVEEIMEKSGEEMPD
LAHVRIILSAENIPNLPGGLAGKNVIEAVYSLNPHRESDGASDAESGSGKL
VEALROMRINHRGNVROLLEMLTSYRLAKVEGESPAEPAATATSSNADAGNPVTMQ
ESHTESGSLAELDSNEDAGTKMSGEKI"
82402..127740
/gene="SLC3A1"
join(82402..82907,87658..87837,88329..88483,92974..93099,
106913..107032,107945..108069,111085..111280,
119528..119695,120777..120893,127141..127740)
/gene="SLC3A1"
join(82478..82907,87658..87837,88329..88483,92974..93099,
106913..107032,107945..108069,111085..111280,
119528..119695,120777..120893,127141..127581)
/gene="SLC3A1"
/note="Homo sapiens solute carrier family 3 (cystine,
dibasic and neutral amino acid transporters, activator of
cystine, dibasic and neutral amino acid transport), member
1 (SLC3A1), mRNA.; H_NH0024105.2
This gene was based on gi(23308570)"
/codon_start=1
/product="unknown"
/protein_id="AA88955.1"
/db_xref="GI:62630210"

gene

mRNA

CDS

/product="unknown"
/protein_id="AA88955.1"
/db_xref="GI:62630210"

WITSFYKSLKDRYGVEDFREVDFPGTMEFENLVAAIHDGKLKLIIDFIPNHTSD
KHIFOLSTRGKTYDYIWHDCIHTENGKTIIPNNLWSVYGNSSWHFDEVRNQCYFH
QWPKQEDNFRNPQVQREIKILRFLWTKGVDFSLDAVKFLAKHLRDEIOVNKT
QIPDTVTOISELYHDTTQVGHDIIVRSFRQTMQDYSTEGRYRPMCTGVAESIDR
TWYGLPFIQADFPFNLSLMDTVSGNSVETIISWNNPKNPKSNWMMIGPQSD
SLRTSLGNQYVNMMLLFTLPDTPITYGEEIGMGNVAANANESVDYNTLRKSP
MOWDINSNAGSEASNTWLPDTSVDYHTVQKTPRSALKYQDLSLHANELLIN
RGWFLHRNDSHVVTRELDGIDRFIVLVNFGSTLNLHNMISGLPAKWRILST
NSADKSKVDTSIGIFLDKGEGLIFEHTKLLRHQTAFRDRCFVSNRACYSVVLNLY
TSC"

misc_feature

93440..93769
/gene="SLC3A1"
/notes="CpG Island (%G=63.6, o/e=0.64, #CpGs=23)"
complement(128218..135958)
/gene="K1AA0436"
complement(join(128218..128387,128769..128842,
129673..129796,130204..130353,133654..133870,
135879..135958))
/gene="K1AA0436"
complement(join(128298..128387,128769..128842,
129673..129796,130204..130353,133654..133870,
135879..135958))
/gene="K1AA0436"
/notes="Homo sapiens putative L-type neutral amino acid
transporter, mRNA [cDNA clone MGC:21932 IMAGE:4389731],
complete cds.; H NH0024105.3
This gene was based on gi(15341987)"
/codon_start=1
/product="unknown"
/protein_id="AA088956.1"
/db_xref="GI:62630211"
/translations="MDLKNFRPERVVDGWLAVHVRGGELGLOWHADRLTK
KLNGLADLRACIKTLHGQFSQSLTLTAFSAGVLAGALNSNPELVRAVTLPEAF
LDVNTWMDTLTLEELFEENWSPSDEKRNKIKRYCPYQNIKPQHPVSHITAYE
NDRVPLKGVTSYTEKEAIAHAHXDTEGYQTPNIILDIOPGNGHVEDSHKKITA
QIKLYEELGIDTSVFEDLKKYLF"

ORIGIN

Query Match 77.6%; Score 19.4; DB 8; Length 169813;
Best Local Similarity 95.2%; Pred. No. 2.1e+02;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 5 AACGTGAATCTCAGGCTGAGA 25

Db 107117 AACGTGATCTCAGGCTGAGA 107097

RESULT 34

AC073487 Homo sapiens 12 BAC RP11-76217 (Robwell Park Cancer Institute Human
BAC Library) complete sequence.
AC073487
AC073487.34 GI:14578058

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE

1 (bases 1 to 54666)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-Osman,F.R., Allen,C.,
Alsbrooke,S.L., Amarantunga,H.C., Are,J.R., Banks,T., Barbara,J.,
Benton,J., Bimage,K., Blankenburg,K., Bonnin,D., Bouck,J.,
Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,
Burck,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,
Carter,M., Cavaros,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
Chen,Z., Chiu,D., Chowdhry,I., Christopoulos,C., Cleveland,C.D.,
Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,
Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,
Elhaj,C., Emerling,S., Escotto,M., Falls,T., Ferraguto,D.,

Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J.,
Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W.,
Gunaratne,P., Hale,S., Hamilton,K., Han,J., Harris,K., Harris,K.,
Hart,M., Havlak,P., Hawes,A., Hernandez,J., Hernandez,O.,
Hodgson,A., Hogue,M., Holloway,C., Hollins,B., Honsi,P.,
Howard,S., Huber,J., Hulyk,S., Hume,J., Ioshikhes,I., Jackson,L.B.,
Jacobson,B., Jia,Y., Johnson,R., Jollivet,S., Joudan,S.,
Karlosson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lee,E., Lewis,L.C.,
Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,
Loulsegh,H., Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
Ma,J., Maheshwari,M., Mapua,P., Marondel,I., Martin,R.,
Martindale,A., Martinez,E., Massey,E., Mawhney,E., McLeod,M.P.,
Meador,M., Mei,G., Merscher,S., Metzker,M., Miller,A., Miner,G.,
Miner,Z., Mitchell,T., Mohabbat,K., Montgomery,K.T., Morgan,M.,
Morris,S., Moser,M., Neal,D., Nelson,D., Newton,J., Newton,N.,
Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S.,
Ogih,M., Okunolu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,
Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,
Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubokan,I., Rolfe,M.,
Ruiz,S., Savary,G., Scherer,S., Scott,G., Shen,H., Shim,C.,
Shooshitari,N., Sisson,I., Sodergren,E., Sonaika,T., Sparks,A.,
Stanley,H., Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A.,
Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B.,
Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D.,
Vinson,R., Wall,R., Wang,S., Ward-Moore,S., Warren,R.,
Washington,C., Watlington,S., Williams,G., Williamson,A.,
Wleczek,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J.,
Zorilla,S., Kucherlapati,R., Weinstein,G. and Gibbs,R.

Direct Submission
Unpublished
2 (bases 1 to 54666)
Worley,K.C.
Direct Submission
Submitted (19-JUN-2000) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 54666)
Worley,K.C.
Direct Submission
Submitted (30-JUN-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
4 (bases 1 to 54666)
Worley,K.C.
Direct Submission
Submitted (03-NOV-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
5 (bases 1 to 54666)
Worley,K.C.
Direct Submission
Submitted (19-MAR-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jun 30, 2001 this sequence version replaced gi:14575758.
INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email
gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the
entire insert of this clone. Overlapping regions of clones are only
sequenced and submitted once, so the sequence for the remainder of
the insert may be found in the record for the adjacent clones.
Overlapping clones are noted at the beginning and end of the
Features listing.

ANNOTATION OF FEATURES:

of SNPs are identified using ePCR (Genome Res. 7:541-550) searches
of a local database that includes entries from dbSTS, GDB, and
local mapping efforts.
Repeats are identified using RepeatMasker (A. Smit and P. Green,
unpublished.) for Human and Mouse sequences.
Genes and Region of sequence similarity are identified by BLAST

(Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL:
<http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

QUALSTAT-REPORT.

Location/Qualifiers

1..54666
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosomes="12"
/clone="RP11-76217"

repeat_region

8..292

/rpt_family="AluSp"

repeat_region

293..320

/rpt_family="C(AAAA)n"

repeat_region

337..454

/rpt_family="AluJb"

repeat_region

456..734

/rpt_family="AluSx"

repeat_region

complement(1548..1602)

/rpt_family="L2"

repeat_region

complement(416..4163)

/rpt_family="MIR"

repeat_region

4164..4445

/rpt_family="AluSx"

repeat_region

complement(4446..4588)

/rpt_family="MIR"

repeat_region

complement(4643..5004)

/rpt_family="L2"

repeat_region

5027..5086

/rpt_family="L2"

repeat_region

complement(5096..5404)

/rpt_family="AluSx"

repeat_region

complement(5738..5799)

/rpt_family="MIR"

repeat_region

complement(6208..6518)

/rpt_family="AluY"

repeat_region

7220..7355

/rpt_family="C-rich"

repeat_region

7583..7801

/rpt_family="CT-rich"

repeat_region

7925..7977

/rpt_family="C(CCG)n"

repeat_region

8037..8076

/rpt_family="C-rich"

repeat_region

8212..8251

/rpt_family="C(GG)n"

repeat_region

10383..10412

/rpt_family="C(GGGGA)n"

repeat_region

10522..10553

/rpt_family="T(CCCC)n"

repeat_region

complement(11756..11898)

/rpt_family="L2"

repeat_region

14938..15072

/rpt_family="C(CAGAGA)n"

STS

/standard_name="21494"
14988..15146
/standard_name="37028"
15260..15281
/rpt_family="AT-rich"
16229..16286
/rpt_family="MIR"
16447..16571
/standard_name="79124"
16559..16661
/standard_name="36508"
16927..17081
/standard_name="87392"
18299..18606
/rpt_family="AluSc"
18612..18788
/rpt_family="FAM"
complement(19482..19685)
/rpt_family="MIR"
21682..21805
/standard_name="73682"
22161..22330
/standard_name="4986"
22219..22335
/standard_name="14903"
23176..23469
/rpt_family="AluSg"
23543..23715
/rpt_family="MSTA"
complement(23717..23966)

Query Match

Best Local Similarity 76.8%; Score 19.2; DB 8; Length 54666;

Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAATCGAATCTCAGGCTGAGA 25

|||||

Db 13288 AAGAACTGAAAATCAGGCTGAGA 13311

RESULT 35

AC026926/c

LOCUS

AC026926

DEFINITION

AC026926

AC026926.1

HTG: HTGS_PHASE0.

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Hominidae; Homo.

REFERENCE

1 (bases 1 to 62880)

Authors

Birren, B., Linton, L., Nusbaum, C. and Lander, E.

Unpublished

JOURNAL

REFERENCE

2 (bases 1 to 62880)

Authors

Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,

Anderson, S., Baldwin, J., Barna, N., Bastien, V., Bedalov, F.,

Boguslavsky, L., Boukhgalter, B., Brown, A., Burkett, G.,

Campopiano, A., Castle, A., Choquel, Y., Colangelo, M., Collins, S.,

Collins, A., Cooke, P., DeArrellano, K., Dewar, K., Diaz, J. S.,

Dodge, S., Domino, M., Doyle, M., Ferreira, P., FitzHugh, W., Gage, D.,

Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,

Grand-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L.,

Howland, J. C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,

Klein, J., LaRocque, K., Lamazares, R., Landers, T., Lehoczy, J.,

Levine, R., Lieu, C., Liu, G., Locke, K., Macdonald, P., Marquis, N.,

McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPheeters, R.,

Meldrum, J., Meneus, L., Minova, T., Miranda, C., Mlenga, V., Morrow, J.,

Murphy, T., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P.,

O'Neil, D., Oliver, T. M., Oliver, J., Peterson, K., Pierre, N.,

Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,

AC026926 62880 bp DNA linear HTG 25-MAR-2000

Homo sapiens clone RP11-658L6, LOW-PASS SEQUENCE SAMPLING.

AC026926.1 GI:7328791

HTG: HTGS_PHASE0.

Homo sapiens (human)

Homo sapiens

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Hominidae; Homo.

REFERENCE

1 (bases 1 to 62880)

Authors

Birren, B., Linton, L., Nusbaum, C. and Lander, E.

Unpublished

JOURNAL

REFERENCE

2 (bases 1 to 62880)

Authors

Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,

Anderson, S., Baldwin, J., Barna, N., Bastien, V., Bedalov, F.,

Boguslavsky, L., Boukhgalter, B., Brown, A., Burkett, G.,

Campopiano, A., Castle, A., Choquel, Y., Colangelo, M., Collins, S.,

Collins, A., Cooke, P., DeArrellano, K., Dewar, K., Diaz, J. S.,

Dodge, S., Domino, M., Doyle, M., Ferreira, P., FitzHugh, W., Gage, D.,

Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,

Grand-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L.,

Howland, J. C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,

Klein, J., LaRocque, K., Lamazares, R., Landers, T., Lehoczy, J.,

Levine, R., Lieu, C., Liu, G., Locke, K., Macdonald, P., Marquis, N.,

McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPheeters, R.,

Meldrum, J., Meneus, L., Minova, T., Miranda, C., Mlenga, V., Morrow, J.,

Murphy, T., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P.,

O'Neil, D., Oliver, T. M., Oliver, J., Peterson, K., Pierre, N.,

Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,

Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,
 Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
 Testaye, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J.,
 Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J.,
 Young, G., Zainoun, J., Zimmer, A. and Zody, M.

TITLE JOURNAL

COMMENT

Direct Submission
 Submitted (25-MAR-2000) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 All repeats were identified using RepeatMasker:

Smit, A. F. A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L7538

Center clone name: 658_L_6

* NOTE: This record contains 69 individual
 * sequencing reads that have not been assembled into
 * contigs. Runs of N are used to separate the reads
 * and the order in which they appear is completely
 * arbitrary. Low-pass sequence sampling is useful for
 * identifying clones that may be gene-rich and allows
 * overlap relationships among clones to be deduced.
 * However, it should not be assumed that this clone
 * will be sequenced to completion. In the event that
 * the record is updated, the accession number will
 * be preserved.

1 767: contig of 767 bp in length
 * 768 867: gap of 100 bp
 * 868 1682: contig of 795 bp in length
 * 1683 1762: gap of 100 bp
 * 1763 2581: contig of 819 bp in length
 * 2582 2681: gap of 100 bp
 * 2682 3470: contig of 789 bp in length
 * 3471 3570: gap of 100 bp
 * 3571 4375: contig of 805 bp in length
 * 4376 4475: gap of 100 bp
 * 4476 5296: contig of 821 bp in length
 * 5297 5396: gap of 100 bp
 * 5397 6170: contig of 774 bp in length
 * 6171 6270: gap of 100 bp
 * 6271 7100: contig of 830 bp in length
 * 7101 7200: gap of 100 bp
 * 7201 8044: contig of 844 bp in length
 * 8045 8144: gap of 100 bp
 * 8145 8947: contig of 803 bp in length
 * 8948 9047: gap of 100 bp
 * 9048 9855: contig of 808 bp in length
 * 9856 9955: gap of 100 bp
 * 9956 10781: contig of 826 bp in length
 * 10782 10881: gap of 100 bp
 * 10882 11718: contig of 837 bp in length
 * 11719 11818: gap of 100 bp
 * 11819 12636: contig of 818 bp in length
 * 12637 12736: gap of 100 bp
 * 12737 13533: contig of 797 bp in length
 * 13534 13633: gap of 100 bp
 * 13634 14458: contig of 823 bp in length
 * 14459 14556: gap of 100 bp
 * 14557 15365: contig of 809 bp in length
 * 15366 15465: gap of 100 bp
 * 15466 16267: contig of 802 bp in length
 * 16268 16367: gap of 100 bp
 * 16368 17173: contig of 806 bp in length
 * 17174 17273: gap of 100 bp
 * 17274 18093: contig of 820 bp in length
 * 18094 18193: gap of 100 bp
 * 18194 19024: contig of 831 bp in length
 * 19025 19124: gap of 100 bp

19125 19927: contig of 803 bp in length
 * 19928 20027: gap of 100 bp
 * 20028 20860: contig of 833 bp in length
 * 20861 20960: gap of 100 bp
 * 20961 21688: contig of 728 bp in length
 * 21689 21788: gap of 100 bp
 * 21789 22620: contig of 832 bp in length
 * 22621 22721: gap of 100 bp
 * 22722 23530: contig of 810 bp in length
 * 23531 23630: gap of 100 bp
 * 23631 24468: contig of 838 bp in length
 * 24469 24568: gap of 100 bp
 * 24569 25366: contig of 798 bp in length
 * 25367 25466: gap of 100 bp
 * 25467 26282: contig of 816 bp in length
 * 26283 26382: gap of 100 bp
 * 26383 27204: contig of 822 bp in length
 * 27205 27304: gap of 100 bp
 * 27305 28131: contig of 827 bp in length
 * 28132 28231: gap of 100 bp
 * 28232 29065: contig of 834 bp in length
 * 29066 29165: gap of 100 bp
 * 29166 30004: contig of 839 bp in length
 * 30005 30104: gap of 100 bp
 * 30105 30941: contig of 837 bp in length
 * 30942 31041: gap of 100 bp
 * 31042 31835: contig of 794 bp in length
 * 31836 31935: gap of 100 bp
 * 31936 32761: contig of 826 bp in length
 * 32762 32861: gap of 100 bp
 * 32862 33694: contig of 833 bp in length
 * 33695 33794: gap of 100 bp
 * 33795 34602: contig of 808 bp in length
 * 34603 34702: gap of 100 bp
 * 34703 35527: contig of 825 bp in length
 * 35528 35627: gap of 100 bp
 * 35628 36442: contig of 815 bp in length
 * 36443 36542: gap of 100 bp
 * 36543 37376: contig of 834 bp in length
 * 37377 37476: gap of 100 bp
 * 37477 38264: contig of 788 bp in length
 * 38265 38364: gap of 100 bp
 * 38365 39135: contig of 771 bp in length
 * 39136 39235: gap of 100 bp
 * 39236 40045: contig of 810 bp in length
 * 40046 40145: gap of 100 bp
 * 40146 40945: contig of 800 bp in length
 * 40946 41045: gap of 100 bp
 * 41046 41762: contig of 717 bp in length
 * 41763 41862: gap of 100 bp
 * 41863 42696: contig of 834 bp in length
 * 42697 42796: gap of 100 bp
 * 42797 43635: contig of 839 bp in length
 * 43636 43735: gap of 100 bp
 * 43736 44630: contig of 795 bp in length
 * 44631 44630: gap of 100 bp
 * 44631 45456: contig of 826 bp in length
 * 45457 45556: gap of 100 bp
 * 45557 46382: contig of 826 bp in length
 * 46383 46482: gap of 100 bp
 * 46483 47297: contig of 815 bp in length
 * 47298 47397: gap of 100 bp
 * 47398 48210: contig of 813 bp in length
 * 48211 48310: gap of 100 bp
 * 48311 49152: contig of 842 bp in length
 * 49153 49252: gap of 100 bp
 * 49253 50063: contig of 811 bp in length
 * 50064 50163: gap of 100 bp
 * 50164 50984: contig of 821 bp in length
 * 50985 51084: gap of 100 bp
 * 51085 51922: contig of 838 bp in length
 * 51923 52022: gap of 100 bp
 * 52023 52841: contig of 819 bp in length

* 52842 52941: gap of 100 bp
* 52942 53770: contig of 829 bp in length
* 53771 53870: gap of 100 bp
* 53871 54648: contig of 778 bp in length
* 54649 54748: gap of 100 bp
* 54749 55542: contig of 794 bp in length
* 55543 55642: gap of 100 bp
* 55643 56466: contig of 824 bp in length
* 56467 56566: gap of 100 bp
* 56567 57401: contig of 835 bp in length
* 57402 57501: gap of 100 bp
* 57502 58319: contig of 818 bp in length
* 58320 58419: gap of 100 bp
* 58420 59234: contig of 815 bp in length
* 59235 59334: gap of 100 bp
* 59335 60165: contig of 831 bp in length
* 60166 60265: gap of 100 bp
* 60266 61081: contig of 816 bp in length
* 61082 61181: gap of 100 bp
* 61182 61969: contig of 788 bp in length
* 61970 62070: contig of 811 bp in length.
* 62070 62880: contig of 811 bp in length.

FEATURES

Location/Qualifiers
Query Match 76.8%; Score 19.2; DB 14; Length 62880;
Best Local Similarity 87.5%; Pred. No. 2.8e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAACTGGAATCTCAGGCTGAG 24
|||||
Db 26104 AAAAACTGATCTCAGCATGAG 26081
|||||

RESULT 36

LOCUS BX255893 99176 bp DNA linear VRT 02-DEC-2004
Zebrafish DNA sequence from clone CH211-272H10 in linkage group 4
DEFINITION Contains the btgl gene for B-cell translocation gene 1, complete sequence.
BX255893

ACCESSION BX255893.5 GI:32260923

VERSION HTG; btgl.

KEYWORDS Danio rerio (zebrafish)

SOURCE

ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 99176)
Dunn, M.

REFERENCE

AUTHORS Direct Submission
TITLE Submitted (02-DEC-2004) Wellcome Trust Sanger Institute, Hinxton,
JOURNAL Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
zf1sh-help@sanger.ac.uk
http://www.sanger.ac.uk/Projects/D_rerio/fags.shtml#dataeight
On Jun 25, 2003 this sequence version replaced gi:32169125.

COMMENT

----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: zf1sh-help@sanger.ac.uk

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:
Em:, EMBL; Sw:, SWISSPROT; Tr:, TrEMBL; Wp:, WORMPEP; Information on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormep/Clone-derived Zebrafish pUC subclones occasionally display inconsistency over the length of mononucleotide A/T runs and conserved TA repeats. Where this is found the longest good quality representation will be submitted.

CH211-272H10 is from a CHORI-211 BAC library
VECTOR: pTARBAC2.1

This sequence was finished as follows unless otherwise noted: all

regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.
Any regions longer than 1kb tagged as misc-feature 'unsure' are part of a tandem repeat of more than 10kb in length where it has not been possible to anchor the base differences between repeat copies. The region has been built up based on the repeat element to match the total size of repeat indicated by restriction digest, but repeat copies may not be in the correct order and the usual finishing criteria may not apply.

FEATURES

source

Location/Qualifiers
1..99176
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/chromosome="4"
/clone="CH211-272H10"
/clone_lib="CHORI-211"
2000
misc_feature
note="Clone left end: CH211-11C20"
23138..23184
misc_feature
note="Single clone region. Assembly confirmed by restriction digest data."
36912..36917
36941
polyA signal
polyA_site
gene
complement(62236..71148)
/gene="btgl"
/locus_tag="CH211-272H10.1-001"
complement(join(62236..65494,70781..71148))
/gene="btgl"
/locus_tag="CH211-272H10.1-001"
/product="B-cell translocation gene 1"
/note="match: cDNAs: Em:BC056691.1
match: ESTs: Em:AW175489.1 Em:AW421052.1 Em:BE201884.1 Em:BE557354.1 Em:BE605378.1 Em:BE605928.1 Em:BE606028.1 Em:BI878310.1 Em:BI878310.1 Em:BI880499.1 Em:BG891842.1 Em:BI325984.1 Em:BI891295.1 Em:BQ92276.1 Em:BI881139.1 Em:BI888289.1 Em:BU710195.1 Em:CA474159.1 Em:BQ092719.1 Em:BQ132066.1 Em:CA474859.1 Em:CA496493.1 Em:CA474442.1 Em:CA474859.1 Em:CA496493.1 Em:CA353813.1 Em:CD283714.1 Em:CD587119.1 Em:CD59529.1 Em:CD605847.1 Em:CD778016.1 Em:CF103301.1 Em:CF924881.1 Em:CF996149.1 Em:CK015896.1 Em:CK016458.1 Em:CK018724.1 Em:CK024547.1 Em:CK236966.1 Em:CK238220.1 Em:CK339087.1 Em:CK354840.1 Em:CK361434.1 Em:CK362051.1 Em:CK397124.1 Em:CK401068.1"
complement(join(65091..65494,70781..70925))
/gene="btgl"
/locus_tag="CH211-272H10.1-001"
/standard_name="OTTDARP0000007966"
/note="match: proteins: Sw:P31607 Sw:P34743 Sw:P53348 Sw:Q63073 Tr:AAH56691"
/codon_start=1
/product="B-cell translocation gene 1"
/protein_id="CAH68953.1"
/db_xref="GI:55251104"
translation="MHILCARGTMKPEINAAVGLSRFLRIKHVNDRLQTFQSOTLQ
DILAQYKHHWFPDRPNKSGRCIRINHKMDPLVGGQGRIGLSIQYLILPSELTL
LWVDPFVYSYRIGEDSGICVLVYBSHPGTNGNPSTTGNISPASSVTQVSAMVESHIS
KEELLVLGRTSPAKPYMMTVSS"
99176
/note="Clone left end: CH211-272H10"

CDS

misc_feature

ORIGIN

Query Match 76.8%; Score 19.2; DB 5; Length 99176;
Best Local Similarity 87.5%; Pred. No. 2.7e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 2 AAAAACTGGAATCTCAGGCTGAGA 25
|||||
Db 70252 AAAAACTGGAATCCAGCTTTAGA 70275
|||||

RESULT 37

AC090678/c

LOCUS

DEFINITION

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

Web site: <http://www.hgsc.bcm.tmc.edu/>

Drafting Center Code: BCM

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: HQ01

Center clone name: RP11-412M2

----- Summary Statistics

Sequencing vector: Plasmid;

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.990329

Consensus quality: 131977 bases at least Q40

Consensus quality: 14820 bases at least Q30

Consensus quality: 150638 bases at least Q20

Estimated insert size: 3481; sum-of-contigs estimation

Quality coverage: 0x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length

* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).

* NOTE: This is a 'working draft' sequence. It currently

* consists of 31 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence.

* as soon as it is available and the accession number will

* be preserved.

* 1 2113: contig of 2113 bp in length

* 2114 2123: gap of unknown length

* 2214 4575: contig of 2362 bp in length

* 4576 4675: gap of unknown length

* 4676 6955: contig of 2180 bp in length

* 6956 9483: gap of unknown length

* 9484 9583: contig of 2528 bp in length

* 9584 11332: contig of 2049 bp in length

* 11333 11732: gap of unknown length

* 11733 13905: contig of 2173 bp in length

* 13906 14005: gap of unknown length

* 14006 16168: contig of 2063 bp in length

* 16169 18363: contig of 2195 bp in length

* 18364 18463: gap of unknown length

* 18464 20771: contig of 2208 bp in length

* 20772 22976: contig of 2205 bp in length

* 22977 23076: gap of unknown length

* 23077 25230: contig of 2154 bp in length

* 25231 28408: contig of 3078 bp in length

* 28409 31188: contig of unknown length

* 31189 31288: gap of unknown length

* 31289 33857: contig of 2569 bp in length

* 33858 33957: gap of unknown length

* 33958 35994: gap of unknown length

* 35994 40174: contig of 4081 bp in length

* 40175 40274: gap of unknown length

* 40275 43333: contig of 3059 bp in length

* 43334 45838: contig of 2405 bp in length

* 45839 45939: gap of unknown length

* 45939 48604: contig of 2666 bp in length

* 48605 48704: gap of unknown length

* 48705 51037: contig of 2333 bp in length

* 51038 51137: gap of unknown length

* 51138 54665: contig of 3528 bp in length

* 54666 54766: gap of unknown length

* 54766 58612: contig of 3847 bp in length

* 58613 58713: gap of unknown length

* 58713 62606: contig of 3894 bp in length

* 62607 62706: gap of unknown length

* 62706 67021: contig of 4315 bp in length

* 67021 67027

REFERENCE

AC090678/c

LOCUS

DEFINITION

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

AC090678

* 67022 67121: gap of unknown length
* 67122 70696: contig of 3575 bp in length
* 70697 70796: gap of unknown length
* 70797 75078: contig of 4282 bp in length
* 75079 75178: gap of unknown length
* 75179 80500: contig of 5322 bp in length
* 80501 80600: gap of unknown length
* 80601 86039: contig of 5439 bp in length
* 86040 86139: gap of unknown length
* 86140 91854: contig of 5715 bp in length
* 91855 91954: gap of unknown length
* 91955 98398: contig of 6444 bp in length
* 98399 98498: gap of unknown length
* 98499 105736: contig of 7238 bp in length.

FEATURES

source

1..105736
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="12"
/clone="RP11-412M2"
2114..2213
/estimated_length=unknown
4576..4675
/estimated_length=unknown
6856..6955
/estimated_length=unknown
9484..9583
/estimated_length=unknown
11633..11732
/estimated_length=unknown
13906..14005
/estimated_length=unknown
16069..16168
/estimated_length=unknown
18364..18463
/estimated_length=unknown
20672..20771
/estimated_length=unknown
22977..23076
/estimated_length=unknown
25231..25330
/estimated_length=unknown
28409..28508
/estimated_length=unknown
31189..31288
/estimated_length=unknown
33858..33957
/estimated_length=unknown
35994..36093
/estimated_length=unknown
40175..40274
/estimated_length=unknown
43334..43433
/estimated_length=unknown
45839..45938
/estimated_length=unknown

Query Match 76.8%; Score 19.2; DB 14; Length 105736;
Best Local Similarity 87.5%; Pred.No.2.7e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAAGTGAATCTCAGGCTGAGA 25

Db 83675 AAGAACTGGAATCAGGCTGAGA 83652

RESULT 38

BX936340/C

LOCUS

BX936340 151328 bp DNA linear HTG 11-AUG-2005

DEFINITION Danio rerio chromosome 12 clone CH211-286P18, WORKING DRAFT

SEQUENCE, 4 unordered pieces.

ACCESSION BX936340

VERSION
KEYWORDS
SOURCE
ORGANISM

BX936340.5 GI:72534258
HTG; HTGS_PHASE1; HTGS_ACTIVEFIN; HTGS_DRAFT; HTGS_FULLTOP.
Danio rerio (zebrafish)
Danio rerio

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.

REFERENCE
1 (bases 1 to 151328)

Almeida,J.

Direct Submission

Submitted (10-AUG-2005) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:

zfish-help@sanger.ac.uk Clone requests:

http://www.sanger.ac.uk/Projects/D_rerio/fags.shtml#datasight

On Aug 12, 2005 this sequence version replaced gi:71533074.

COMMENT

----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: zfish-help@sanger.ac.uk
----- Project Information
Center project name: zc286P18
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; 100% of reads
Consensus quality: 149941 bases at least Q40
Consensus quality: 150354 bases at least Q30
Consensus quality: 150523 bases at least Q20
Insert size: 151028; sum-of-contigs
Quality coverage: 9.14x in Q20 bases; sum-of-contigs Quality
coverage: 8.46x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 14418: contig of 14418 bp in length
* 14419 14518: gap of 100 bp
* 14519 47576: contig of 33058 bp in length
* 47577 47676: gap of 100 bp
* 47677 149149: contig of 101473 bp in length
* 149150 149249: gap of 100 bp
* 149250 151328: contig of 2079 bp in length.

FEATURES

source

1..151328
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/chromosome="12"
/clone="CH211-286P18"
/clone_lib="CHORI-211"

misc_feature

1..14418
/note="assembly_fragment:00760
fragment_chain:1"

misc_feature

14519..47576
/note="assembly_fragment:02430
fragment_chain:1"

misc_feature

47677..149149
/note="assembly_fragment:02607
fragment_chain:1"

misc_feature

149250..151328
/note="assembly_fragment:02573"

ORIGIN

Query Match 76.8%; Score 19.2; DB 14; Length 151328;

Best Local Similarity 87.5%; Pred.No.2.6e+02;

Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

repeat_region	/rpt_family=" (TCCC)n" 16482. .16657 /rpt_family=" (TTTC)n" complement(16675. .16754) /rpt_family="ID4_ " 17131. .17233 /rpt_family="WTE " 17994. .18569 /rpt_family="MTB2 " complement(19676. .19951) /rpt_family="B4A " 20123. .20493 /rpt_family="ORR1D " complement(20976. .21179) /rpt_family="B3 " 21351. .21371 /rpt_family="AT rich " complement(21917. .22093) /rpt_family="B4A " complement(22154. .22282) /rpt_family="L2 " complement(23052. .23145) /rpt_family="L2 " 23823. .23907 /rpt_family="CT rich " complement(24990. .25155) /rpt_family="ID_B1 " 25258. .25437 /rpt_family=" (TTTC)n " complement(25439. .25590) /rpt_family="Lx6 " 26606. .26651 /rpt_family=" (TG)n " complement(26931. .27291) /rpt_family="ORR1D " complement(27369. .27406) /rpt_family="RMER1A " complement(27435. .27671) /rpt_family="RMER1C " 27645. .27749 /rpt_family="RMER1C " 27839. .27937 /rpt_family="MIR " complement(29389. .30572) /rpt_family="RMER12 " 30984. .31173 /rpt_family="B2_Mm2 " 31176. .31298 /rpt_family=" (CGA)n " 31946. .32172 /rpt_family="MTB_Mm " 32176. .32197 /rpt_family="AT rich " 32594. .32656 /rpt_family="MTB_Mm " 32728. .32917 /rpt_family="B3 " 33223. .33317 /rpt_family="B1F " 33286. .33467 /standard_name="VHAa67g10.seq"	76.8%; Score 19.2; DB 9; Best Local Similarity 87.5%; Pred. No. 2.6e+02; Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;	Query Match 2 AAAAAGTGAATCTCAGGCTGAGA 25 83238 AAAAAGTGAATCTCAGGCTGAGA 83215 	ST5
repeat_region	LOCUS AC104037 152267 bp DNA linear PRI 12-MAR-2000 DEFINITION Homo sapiens chromosome 8, clone RP11-692P18, complete sequence. ACCESSION AC104037 VERSION AC104037.3 GI:19352312 KEYWORDS HTG. SOURCE Homo sapiens (human) ORGANISM Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 152267) Birren,B., Linton,L., Nusbaum,C. and Lander,E. Homo sapiens chromosome 8, clone RP11-692P18 Unpublished 2 (bases 1 to 152267) Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kanat,A., Karatas,A., Kells,C., Labrecque,K., Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G., MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Murphy,T., Naylor,J., Nguyen,C., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupack,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M. Direct Submission Submitted (01-DEC-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA 3 (bases 1 to 152267) Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kanat,A., Karatas,A., Kells,C., Labrecque,K., Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G., MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Murphy,T., Naylor,J., Nguyen,C., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupack,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M. Direct Submission Submitted (12-MAR-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA On Mar 12, 2002 this sequence version replaced gi:18377276. All repeats were identified using RepeatMasker: Smit, A.F.A. & Green, P. (1996-1997) http://ftp.genome.washington.edu/RM/RepeatMasker.html ----- Genome Center Center: Whitehead Institute/ MIT Center for Genome Research Genome Center code: WIBR			

Web site: <http://www-seq.wi.mit.edu>
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L22011
 Center clone name: 692_P_18

FEATURES

Source

```

1. .152267
   /organism="Homo sapiens"
   /mol_type="genomic DNA"
   /db_xref="taxon:9606"
   /chromosome="8"
   /map="8"
   /clone="RP11-692P18"
   /clone_lib="RPC1-11 Human Male BAC"
2. .440
   /rpt_family="L1MD1"
35. .846
   /rpt_family="L1MD1"
   /complement(2002. .2456)
   /rpt_family="MER115"
3150. .3170
   /rpt_family="AT_rich"
3204. .6714
   /rpt_family="L1PA10"
6767. .6982
   /rpt_family="L2"
   /complement(6994. .7071)
   /rpt_family="Tigger7"
7072. .7179
   /rpt_family="MER47A"
   /complement(7181. .7309)
   /rpt_family="L1PA16"
7319. .7881
   /rpt_family="L1PB2"
   /complement(7892. .9413)
   /rpt_family="L1PA16"
9414. .9662
   /rpt_family="MER47A"
   /complement(9677. .9938)
   /rpt_family="Tigger7"
9942. .9973
   /rpt_family="L2"
   /complement(9969. .10388)
   /rpt_family="L2"
   /complement(10394. .10453)
   /rpt_family="MIR"
   /complement(10653. .10957)
   /rpt_family="AluSx"
11200. .11226
   /rpt_family="(A)n"
11402. .11666
   /rpt_family="LTR16A1"
14862. .14892
   /rpt_family="AT_rich"
   /complement(15052. .15120)
   /rpt_family="MLTIF1"
   /complement(15123. .15398)
   /rpt_family="AluSx"
15445. .15688
   /rpt_family="Aluub"
15698. .15876
   /rpt_family="MIR"
16774. .17292
   /rpt_family="MLTIF1"
17493. .19099
   /rpt_family="L1PA13"
19620. .20096
   /rpt_family="L2"
20267. .20552
   /rpt_family="Aluub"
20557. .20991
   /rpt_family="L2"

```

```

repeat_region 21007. .21277
                /rpt_family="AluSg"
repeat_region complement(21924. .22230)
                /rpt_family="AluSg"
repeat_region complement(23277. .23569)
                /rpt_family="AluSx"
repeat_region complement(23872. .23913)
                /rpt_family="MIR3"
repeat_region 23956. .24031
                /rpt_family="L1MCS"
repeat_region complement(24752. .25065)
                /rpt_family="L2"
repeat_region complement(25942. .26173)
                /rpt_family="AluJo"
repeat_region complement(27746. .27906)
                /rpt_family="MER112"
repeat_region 27919. .28057
                /rpt_family="MIR"
repeat_region complement(28222. .28337)
                /rpt_family="L2"
repeat_region complement(29041. .29436)
                /rpt_family="L1PA5"
repeat_region 29442. .29626
                /rpt_family="L1PA5"
repeat_region 30000. .30208
                /rpt_family="MER58A"
repeat_region 30932. .30996
                /rpt_family="GA-rich"
repeat_region 31046. .31399
                /rpt_family="MLT1C"
repeat_region complement(31424. .32306)
                /rpt_family="L1MA7"
repeat_region complement(32307. .33165)
                /rpt_family="L1"
repeat_region complement(33166. .33454)
                /rpt_family="AluSx"
repeat_region complement(33455. .33804)
                /rpt_family="L1"
repeat_region 33805. .34087
                /rpt_family="AluJb"
repeat_region complement(34088. .34296)
                /rpt_family="L1"
repeat_region complement(34334. .35294)
                /rpt_family="L1"
repeat_region complement(35305. .35452)
                /rpt_family="MLT1B"
repeat_region complement(35556. .36179)
                /rpt_family="L1PA4"
repeat_region complement(36188. .36459)
                /rpt_family="L1M2"
repeat_region complement(37672. .37816)
                /rpt_family="L1M2"
repeat_region 37830. .37940
                /rpt_family="MLT1C"

```

Query Match 76.8%; Score 19.2; DB 8; Length 152267;
 Best Local Similarity 87.5%; Pred. No. 2.6e+02;
 Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAAAGTGAATCTCAGGCTGAG 24
 |||||
 Db 43430 AAGAAGTGAATCTCAGTCTGAG 43407

RESULT 41

AL671979/c 154170 bp DNA linear ROD 05-APR-2002
 LOCUS Mouse DNA sequence from clone RP23-314B7 on chromosome X, complete
 DEFINITION
 accession AL671979 sequence.
 version AL671979.5 GI:20068746
 keywords HTG.
 source Mus musculus (house mouse)

ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1
AUTHORS Matthews, L.
TITLE Direct Submission
JOURNAL Submitted (04-APR-2002) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
COMMENT On Apr 8, 2002 this sequence version replaced gi:19743946. During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above. This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep/ RP23-314B7 is from the RPI-23 Mouse PAC Library constructed by the group of Pieter de Jong. For further details see http://www.chori.org/bacpac/home.htm VECTOR: pBACe3.6.

FEATURES
source
1. .154170
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="X"
/clone="RP23-314B7"
/clone_lib="RPCI-23"

ORIGIN
Query Match 76.8%; Score 19.2; DB 9; Length 154170;
Best Local Similarity 87.5%; Pred. No. 2.6e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 2 AAAAACTGGAATCTCAGGCTGAGA 25
Db 8104 AAAAACTGGAATCTTAGGCAGAGA 8081
|||||
AC132332 155268 bp DNA linear ROD 29-MAY-2004
Mus musculus BAC clone RP24-222H7 from chromosome 10, complete sequence.
AC132332 GI:47825139
AC132332.2
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 155268)
AUTHORS Tomlinson, C., Haglund, K., Bielicki, L. and Meyer, R.
TITLE The sequence of Mus musculus BAC clone RP24-222H7
JOURNAL Unpublished (2001)
REFERENCE 2 (bases 1 to 155268)
AUTHORS McPherson, J. D. and Waterston, R. H.
TITLE Direct Submission
JOURNAL Submitted (03-SEP-2002) Genome Sequencing Center, 4444 Forest Park

REFERENCE 3 (bases 1 to 155268)
AUTHORS Wilson, R. K.
TITLE Direct Submission
JOURNAL Submitted (29-MAY-2004) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
COMMENT On May 29, 2004 this sequence version replaced gi:22657812. ----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu
Contact: submissions@watson.wustl.edu
----- Summary Statistics
Center project name: M_BB0222H07

NOTICE:
This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. Wes Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see http://genome.wustl.edu

SOURCE INFORMATION:
The RPCI-24 BAC Library has been constructed by Pieter de Jong and coworkers (http://www.chori.org) from male C57BL/6J mouse spleen and/or brain genomic DNA. The clone and detailed information can be obtained from Pieter de Jong and coworkers at http://www.chori.org

NEIGHBORING SEQUENCE INFORMATION:
This sequence is the entire insert of the clone.
Location/Qualifiers
1. .155268
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="10"
/map="10"
/clone="RP24-222H7"
/clone_lib="RPCI-24"
279. .847
/rpt_family="L1"
repeat_region
1834. .2118
/rpt_family="ERV1"
repeat_region
2133. .2555
/rpt_family="L1"
repeat_region
2600. .2744
/rpt_family="ERV1"
repeat_region
3461. .3904
/rpt_family="ACHobo"
repeat_region
4150. .4399
/rpt_family="MalR"
misc_feature
5940. .6781
/note="Sequence derived from PCR product of project DNA."
repeat_region
6067. .6175
/rpt_family="L2"
repeat_region
6348. .6642
/rpt_family="B4"
repeat_region
6803. .7304
/rpt_family="L1"
repeat_region
9443. .9667
/rpt_family="MER1_type"
repeat_region
9861. .10148

```
repeat_region /rpt family="L1"
repeat_region 10149. .10371
repeat_region /rpt family="MER1_type"
repeat_region 10372. .10451
repeat_region /rpt family="L1"
repeat_region 10554. .10668
repeat_region /rpt family="L1"
repeat_region 10666. .10985
repeat_region /rpt family="L1"
repeat_region 11521. .11701
repeat_region /rpt family="B2"
repeat_region 12326. .12382
repeat_region /rpt family="L1"
repeat_region 12385. .12477
repeat_region /rpt family="L1"
repeat_region 12721. .12859
repeat_region /rpt family="Alu"
repeat_region 13133. .13248
repeat_region /rpt family="L1"
repeat_region 14350. .14424
repeat_region /rpt family="ID"
repeat_region 14461. .14802
repeat_region /rpt family="ERVk"
repeat_region 15333. .15647
repeat_region /rpt family="ERVl"
repeat_region 16395. .16565
repeat_region /rpt family="MER1_type"
repeat_region 17745. .17857
repeat_region /rpt family="L1"
repeat_region 17858. .18024
repeat_region /rpt family="MER1_type"
repeat_region 18788. .19010
repeat_region /rpt family="B4"
repeat_region 19646. .19889
repeat_region /rpt family="L1"
repeat_region 20866. .21180
repeat_region /rpt family="L1"
repeat_region 21233. .21325
repeat_region /rpt family="L1"
repeat_region 22395. .22602
repeat_region /rpt family="MER1_type"
repeat_region 23561. .23611
repeat_region /rpt family="ERV1"
repeat_region 23752. .23883
repeat_region /rpt family="B4"
repeat_region 23936. .24123
repeat_region /rpt family="ERV1"
repeat_region 24565. .24632
repeat_region /rpt family="ERVl"
repeat_region 24985. .25114
repeat_region /rpt family="ERVl"
repeat_region 25279. .25584
repeat_region /rpt family="B4"
repeat_region 27940. .28070
repeat_region /rpt family="B2"
repeat_region 29585. .29941
repeat_region /rpt family="MaLR"
repeat_region 30046. .30171
repeat_region /rpt family="L1"
repeat_region 31130. .31340
repeat_region /rpt family="B2"
repeat_region 31567. .31624
repeat_region /rpt family="ERV1"
repeat_region 31625. .31757
repeat_region /rpt family="Alu"
repeat_region 31847. .31930
repeat_region /rpt family="L1"
repeat_region 34010. .34213
repeat_region /rpt family="B2"
repeat_region 34530. .34645
repeat_region /rpt family="Alu"
repeat_region 34813. .34932
repeat_region /rpt family="Alu"
```

```
repeat_region 35001. .35428
repeat_region /rpt family="MaLR"
repeat_region 35519. .35874
repeat_region /rpt family="MaLR"
repeat_region 37067. .37511
repeat_region /rpt family="MaLR"
repeat_region 38737. .38875
repeat_region /rpt family="B4"
repeat_region 41238. .41514
repeat_region /rpt family="ERV1"
repeat_region 42758. .43087
repeat_region /rpt family="MaLR"
repeat_region 43088. .43195
repeat_region /rpt family="MaLR"
repeat_region 43212. .43331
repeat_region /rpt family="L1"
repeat_region 45072. .45495
repeat_region /rpt family="MaLR"
repeat_region 45560. .45691
repeat_region /rpt family="Alu"
repeat_region 45823. .46006
repeat_region /rpt family="B2"
repeat_region 46221. .46618
repeat_region /rpt family="MaLR"
repeat_region 47277. .47344
repeat_region /rpt family="ID"
repeat_region 47782. .48147
repeat_region /rpt family="L1"
repeat_region 49447. .49505
repeat_region /rpt family="L1"
repeat_region 49526. .49709
repeat_region /rpt family="L1"
repeat_region 50291. .50443
repeat_region /rpt family="MaLR"
repeat_region 51494. .51898
```

Query Match 76.8%; Score 19.2; DB 9; Length 155268;
Best Local Similarity 87.5%; Pred. No. 2.6e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 2 AAAAACTGGAATCTCAGGCTGAGA 25
|||||
Db 33162 AAAAACTGGAATCTCAGGCTGAGA 33139
|||||

RESULT 43
AF307158/c

LOCUS AF307158 155669 bp DNA linear HTG 26-JUL-2002
DEFINITION Homo sapiens chromosome 8 clone RP11-117P17 map 8p12, WORKING DRAFT
SEQUENCE, 8 unordered pieces.

ACCESSION AF307158.2 GI:14389408
VERSION AF307158.2
KEYWORDS HTGS_PHASE1; HTGS_DRAFT; HTGS_CANCELLED.
SOURCE Homo sapiens (human)

ORGANISM
Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.

REFERENCE 1 (bases 1 to 155669)

AUTHORS Schilthabel, M.B., Baumgart, C., Blechschmidt, K., Dette, M., Jahn, N.,
Lehmann, R., Menzel, U., Polley, A., Reichwald, K., Schudy, A.,
Siddiqui, R., Taudien, S., Wen, G., Rosenthal, A. and Platzner, M.

Chromosome 8 genomic sequence

Unpublished

REFERENCE 2 (bases 1 to 155669)

AUTHORS Genome Sequencing Center Jena.

Direct Submission

TITLE Submitted (21-SEP-2000) Genome Analysis, Institute of Molecular

JOURNAL Biotechnology, Beutenbergstr. 11, Jena 07745, Germany

On Jun 13, 2001 this sequence version replaced gi:10334826.

COMMENT ----- Genome Center

Center: Institute of Molecular Biotechnology

Center code: IMB

and/or brain genomic DNA. The clone and detailed information can be obtained from Pieter de Jong and coworkers at <http://www.chori.org>

NEIGHBORING SEQUENCE INFORMATION:

This sequence is the entire insert of the clone. This clone is overlapped by AC140243.

FEATURES

source

1. .162528

Location/Qualifiers

/organism="Mus musculus"

/mol_type="genomic DNA"

/db_xref="taxon:10090"

/chromosome="9"

/map="g"

/clone="RP24-296C19"

/clone_lib="RPC1-24"

932. .1076

/rpt_family="Alu"

1787. .1971

/rpt_family="MIR"

2617. .2756

/rpt_family="MIR"

5649. .5775

/rpt_family="Alu"

5776. .5857

/rpt_family="B4"

7268. .7672

/rpt_family="RMR17C"

8436. .8519

/rpt_family="MIR"

8689. .8757

/rpt_family="L2"

8817. .8949

/rpt_family="Alu"

9301. .9660

/rpt_family="MaLR"

10124. .10303

/rpt_family="B2"

10355. .10496

/rpt_family="B4"

10571. .10718

/rpt_family="Alu"

10779. .10908

/rpt_family="MaLR"

10906. .11027

/rpt_family="B2"

11070. .11222

/rpt_family="Alu"

11789. .12045

/rpt_family="B4"

12241. .12358

/rpt_family="B2"

12377. .12536

/rpt_family="B4"

12593. .12650

/rpt_family="L1"

12655. .12800

/rpt_family="Alu"

12806. .12922

/rpt_family="L1"

13137. .13281

/rpt_family="Alu"

13289. .13360

/rpt_family="ID"

13532. .13720

/rpt_family="B2"

14411. .14550

/rpt_family="Alu"

17393. .17539

/rpt_family="Alu"

17784. .17977

/rpt_family="B2"

18370. .18480

/rpt_family="B4"

repeat_region

repeat_region

tRNA

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

Query Match

76.8%; Score 19.2; DB 9; Length 162528;

Best Local Similarity 87.5%; Pred. No. 2.6e+02;

Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAACTGAATCTCAGGCTGAG 24

Db 15606 AAAAACTGAATCTCAGGCTGAG 15583

RESULT 45
AC119551
LOCUS
DEFINITION
AC119551
Rattus norvegicus clone CH230-412B6, *** SEQUENCING IN PROGRESS ***
AC119551
AC119551.5 GI:25008679
HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_ENRICHED.
Rattus norvegicus (Norway rat)
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Rattus;
1 (bases 1 to 164792)
Muzny, D.Marie, Metzker, M.Lee, Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowals, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, X., London, P., Longacre, S., Lopez, J., Lorusu, H., Loulseghe, H., Lozado, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, W., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwankwelen, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L., Puaro, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajs, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steinle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villalana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstein, G. and Gibbs, R.A.
Direct Submission
Unpublished
2 (bases 1 to 164792)
Worley, K.C.
Direct Submission
TITLE
JOURNAL
AUTHORS
Submitted (28-APR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
Rat Genome Sequencing Consortium.
3 (bases 1 to 164792)
REFERENCE
AUTHORS

TITLE
JOURNAL
COMMENT
Direct Submission
Submitted (15-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Nov 15, 2002 this sequence version replaced gi:23808213.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GWFF
Center clone name: CH230-412B6
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 152401 bases at least Q40
Consensus quality: 154698 bases at least Q30
Consensus quality: 155987 bases at least Q20
Estimated insert size: 154442; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* the accession number will be 164792 bp in length.
*
Location/Qualifiers
1. 164792
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-412B6"
1. 1456
/note="wgs_end_extension
clone_end:T7"
5329..6087
/note="clone_boundary
clone_end:T7"
site:
end_sequence:BZ142077"
complement(159147..160072)
/note="clone_boundary
clone_end:Sp6"
site:
end_sequence:BZ142078"
161480..162593
/note="wgs_end_extension
clone_end:Sp6"
162644..164792
/note="wgs_end_extension
clone_end:Sp6"
misc_feature
misc_feature
misc_feature
misc_feature
misc_feature
ORIGIN
Query Match 76.8%; Score 19.2; DB 14; Length 164792;

Best Local Similarity 87.5%; Pred. No. 2.6e+02; Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 AAAAACTGGATCTCAGGCTGAG 24
 ||||| ||||| ||||| ||||| |||||
 DB 15404 AAAAAATAGATCTGAGGCTGAG 15427

RESULT 46
 AC132235/c
 LOCUS AC132235 180462 bp DNA linear ROD 15-MAY-2004
 DEFINITION Mus musculus BAC clone RP24-497H15 from chromosome 17, complete
 sequence.

ACCESSION AC132235 GI:46559523
 VERSION AC132235.3
 KEYWORDS HTG.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 180462)
 VanBrunt, A., Cotton, M., Van Brunt, A., Kozlowski, A. and Haakenson, W.
 The sequence of Mus musculus BAC clone RP24-497H15
 Unpublished (2001)

REFERENCE 2 (bases 1 to 180462)
 McPherson, J.D. and Waterston, R.H.
 Direct Submission
 Submitted (03-SEP-2002) Genome Sequencing Center, 4444 Forest Park
 Parkway, St. Louis, MO 63108, USA

REFERENCE 3 (bases 1 to 180462)
 Wilson, R.K.
 Direct Submission
 Submitted (05-MAR-2004) Genome Sequencing Center, 4444 Forest Park
 Parkway, St. Louis, MO 63108, USA

REFERENCE 4 (bases 1 to 180462)
 Wilson, R.K.
 Direct Submission
 Submitted (24-APR-2004) Genome Sequencing Center, 4444 Forest Park
 Parkway, St. Louis, MO 63108, USA

REFERENCE 5 (bases 1 to 180462)
 Wilson, R.K.
 Direct Submission
 Submitted (15-MAY-2004) Genome Sequencing Center, 4444 Forest Park
 Parkway, St. Louis, MO 63108, USA

On Apr 24, 2004 this sequence version replaced gi:45120391.
 ----- Genome Center

Center: Washington University Genome Sequencing Center

Center code: WUGSC

Web site: <http://genome.wustl.edu>

Contact: submissions@watson.wustl.edu

----- Summary Statistics

Center project name: M_BB0497H15

NOTICE:

This sequence was finished as follows unless otherwise noted:
 all regions were double stranded, sequenced with an alternate
 chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by at least
 one plasmid subclone, fosmid clone or direct clone walk sequence.
 Sequence from the Mouse Genome Sequencing Consortium whole genome
 shotgun may have been used to obtain the consensus sequence; and
 the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. Wes Warren,
 Department of Genetics, Washington University, St. Louis MO. For
 additional information about the map position of this sequence, see
<http://genome.wustl.edu>

SOURCE INFORMATION:

The RPCI-24 BAC Library has been constructed by Pieter de Jong and
 coworkers (<http://www.chori.org>) from male C57BL/6J mouse spleen
 and/or brain genomic DNA. The clone and detailed information can be
 obtained from Pieter de Jong and coworkers at <http://www.chori.org>

NEIGHBORING SEQUENCE INFORMATION:

This sequence is the entire insert of the clone. This clone is
 overlapped by AC133946 and AL592112.

FEATURES	Location/Qualifiers
source	1..180462 /organism="Mus musculus" /mol_type="genomic DNA" /db_xref="taxon:10090" /chromosome="17" /map="17" /clone="RP24-497H15" /clone_lib="RPCI-24"
repeat_region	2012..2086 /rpt_family="MIR"
repeat_region	4862..5007 /rpt_family="B2"
repeat_region	5026..5102 /rpt_family="Alu"
repeat_region	5156..5427 /rpt_family="B4"
repeat_region	7815..7914 /rpt_family="MER1_type"
repeat_region	8205..8351 /rpt_family="Alu"
repeat_region	9343..9404 /rpt_family="ERV1"
repeat_region	9657..9838 /rpt_family="ERV1"
repeat_region	10281..10453 /rpt_family="B2"
repeat_region	11584..11807 /rpt_family="ERVK"
repeat_region	12089..12544 /rpt_family="L1"
repeat_region	12621..12693 /rpt_family="MaLR"
repeat_region	13372..13642 /rpt_family="MaLR"
repeat_region	13643..14039 /rpt_family="MaLR"
repeat_region	14040..14155 /rpt_family="MaLR"
repeat_region	14174..14306 /rpt_family="MIR"
repeat_region	14635..14982 /rpt_family="MaLR"
repeat_region	14983..15321 /rpt_family="MaLR"
repeat_region	15406..15530 /rpt_family="Alu"
repeat_region	16695..16846 /rpt_family="L1"
repeat_region	16982..17188 /rpt_family="MaLR"
repeat_region	19686..19783 /rpt_family="Alu"
repeat_region	20689..21013 /rpt_family="MaLR"
repeat_region	21034..21232 /rpt_family="B4"
repeat_region	21355..21657 /rpt_family="B4"
repeat_region	21897..22051 /rpt_family="Alu"
repeat_region	22482..22598 /rpt_family="Alu"
repeat_region	22549..22616

```
repeat_region /rpt_family="B4"
23095. .23166
repeat_region /rpt_family="Alu"
23167. .23314
repeat_region /rpt_family="Alu"
23315. .23354
repeat_region /rpt_family="Alu"
23363. .23523
repeat_region /rpt_family="B4"
26232. .27488
repeat_region /rpt_family="ERVK"
27489. .27886
repeat_region /rpt_family="ERVK"
27887. .27935
repeat_region /rpt_family="B4"
28468. .28614
repeat_region /rpt_family="Alu"
28795. .28934
repeat_region /rpt_family="Alu"
30112. .30175
repeat_region /rpt_family="B2"
30193. .30526
repeat_region /rpt_family="MaLR"
30572. .30683
repeat_region /rpt_family="B2"
31165. .31186
repeat_region /rpt_family="B4"
31187. .31256
repeat_region /rpt_family="ID"
31257. .31361
repeat_region /rpt_family="B4"
34487. .34555
repeat_region /rpt_family="B4"
34561. .34705
repeat_region /rpt_family="Alu"
35095. .35294
repeat_region /rpt_family="ERVK"
36059. .36320
repeat_region /rpt_family="ERVK"
36893. .37050
repeat_region /rpt_family="B4"
37587. .37653
repeat_region /rpt_family="ERV1"
38313. .38542
repeat_region /rpt_family="B4"
38636. .38698
repeat_region /rpt_family="ERV1"
38751. .39053
repeat_region /rpt_family="MaLR"
40128. .40308
repeat_region /rpt_family="L2"
40319. .40385
repeat_region /rpt_family="B4"
40401. .40526
repeat_region /rpt_family="Alu"
40646. .40706
repeat_region /rpt_family="L2"
41898. .42130
repeat_region /rpt_family="MaLR"
42183. .42380
repeat_region /rpt_family="B2"
44039. .44238
repeat_region /rpt_family="B2"
44805. .44904
```

```
Query Match 76.8%; Score 19.2; DB 9; Length 180462;
Best Local Similarity 87.5%; Pred. No. 2.6e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAACTGGAACTCAGGCTGAG 24
||||| ||||||| |||||
Db 29952 AAAAAATGGAACTCAGATGAG 29929
```

RESULT 47
AC130713
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

AC130713 192020 bp DNA linear ROD 15-MAY-2004
Mus musculus BAC clone RP23-316K10 from chromosome 15, complete
sequence.
AC130713
AC130713.4 GI:45120394
HTG.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.
1 (bases 1 to 192020)
Goyea, E., Tomlinson, C., Haakenson, W. and Meyer, R.
The sequence of Mus musculus BAC clone RP23-316K10
Unpublished (2001)
2 (bases 1 to 192020)
McPherson, J.D. and Waterston, R.H.
Direct Submission
Submitted (13-AUG-2002) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
3 (bases 1 to 192020)
McPherson, J.D. and Waterston, R.H.
Direct Submission
Submitted (14-MAY-2003) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
4 (bases 1 to 192020)
Wilson, R.K.
Direct Submission
Submitted (05-MAR-2004) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
5 (bases 1 to 192020)
Wilson, R.K.
Direct Submission
Submitted (15-MAY-2004) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
On Mar 5, 2004 this sequence version replaced gi:30698688.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu>
Contact: submissions@watson.wustl.edu
----- Summary Statistics

Center project name: M_BA0316K10

NOTICE:

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone, fosmid clone or direct clone walk sequence.
Sequence from the Mouse Genome Sequencing Consortium whole genome
shotgun may have been used to obtain the consensus sequence; and
the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. Wes Warren,
Department of Genetics, Washington University, St. Louis MO. For
additional information about the map position of this sequence, see
<http://genome.wustl.edu>

SOURCE INFORMATION:

The RPCT-23 BAC library has been constructed by Kazutoyo Osegawa
and Minako Tatenno in the laboratory of Pieter de Jong
(<http://www.chori.org>) from female C57BL/6J mouse kidney and/or
brain genomic DNA. The clone and detailed information can be
obtained from Research Genetics, Inc. (<http://www.resgen.com>) or
Pieter de Jong and coworkers at <http://www.chori.org>

NEIGHBORING SEQUENCE INFORMATION:

This sequence is the entire insert of the clone. This clone is overlapped by AC116950.

FEATURES

source Location/Qualifiers

1..192020
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="15"
/map="15"

/clone="RP23-316K10"
/clone_lib="RPCI-23"

unsure 17082..17085
/note="Sequence derived from one plasmid subclone."

misc_feature 17418..17861
/notes="Sequence derived from PCR product of project DNA."

unsure 130749..131038
/note="Unresolved simple sequence repeat."

ORIGIN

Query Match 76.8%; Score 19.2; DB 9; Length 192020;
Best Local Similarity 87.5%; Pred. No. 2.6e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAACTGGAATCTAGGCTGAG 24

Db 48721 AAAAACTGGAATCTAGGCTGAG 48744

RESULT 48

AC114552/c
LOCUS Mus musculus chromosome 12, clone RP23-349L22, complete sequence.
DEFINITION AC114552
ACCESSION AC114552
VERSION AC114552.16 GI:45642942

KEYWORDS HTG.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidae; Murinae; Mus.
1 (bases 1 to 197585)
Birren,B., Nusbaum,C. and Lander,E.

REFERENCE Mus musculus chromosome 12, clone RP23-349L22

Unpublished

2 (bases 1 to 197585)

AUTHORS

Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Gord,S., Govette,M., Graham,L., Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., LaRoque,K., Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Lindblad-Toh,K., Liu,G., MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schuback,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Straus,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission

Submitted (10-MAR-2002) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

3 (bases 1 to 197585)

AUTHORS

Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N., Anderson,M., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y., Collymore,A., Cooke,P., Corum,B., Dearellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,X., Lui,A., Mabbitt,R., Maclean,C., Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schuback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission

Submitted (12-FEB-2004) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

4 (bases 1 to 197585)

REFERENCE

Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N., Anderson,M., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y., Collymore,A., Cooke,P., Corum,B., Dearellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,X., Lui,A., Mabbitt,R., Maclean,C., Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schuback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission

Submitted (23-MAR-2004) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

On Mar 23, 2004 this sequence version replaced gi:42539087.

All repeats were identified using RepeatMasker:

Smith, A.F.A. & Green, P. (1996-1997)

<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L22677

Center clone name: 349_L_22

FEATURES

source

Location/Qualifiers
1..197585
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="12"
/map="12"
/clone="RP23-349L22"
/clone_lib="RPCI-23 Female Mouse BAC"
1..6845
/notes="wgs_end_extension
clone_end:SP6"
complement(2440..2644)
repeat_region

misc_feature

1..6845

/notes="wgs_end_extension

clone_end:SP6"

complement(2440..2644)

```
unsure      /rpt_family="B3"
2892_.3010
/note="single clone coverage"
repeat_region /rpt_family=" (TGGGGG)n"
4442_.4475
repeat_region /rpt_family=" (TGGGGG)n"
5059_.5206
repeat_region /rpt_family="MIR"
complement(5391_.5533)
/rpt_family="B1F"
6481_.6511
repeat_region /rpt_family=" (TTGG)n"
complement(6524_.6624)
/rpt_family="RSIN1"
6846_.6851
misc_feature /note="clone boundary
clone end:SP6
site:EcoRI"
repeat_region 7510_.7549
/rpt_family=" (TG)n"
repeat_region 7980_.8352
/rpt_family="MLT1H1"
complement(8694_.8753)
/rpt_family="L2"
9133_.9473
repeat_region /rpt_family="L1_MM"
9838_.9993
repeat_region /rpt_family=" (TCTCCC)n"
9998_.10066
/rpt_family=" (TCTG)n"
complement(10161_.10165)
/note="30 qual SINGL region"
11431_.11597
repeat_region /rpt_family="ORR1A3"
11900_.11991
repeat_region /rpt_family="MER20"
12010_.12054
repeat_region /rpt_family=" (TA)n"
12516_.12537
repeat_region /rpt_family="AT_rich"
12576_.12779
repeat_region /rpt_family="Lx8"
13023_.13082
repeat_region /rpt_family="Lx8"
13124_.13258
repeat_region /rpt_family="B1_MM"
13259_.13477
repeat_region /rpt_family=" (GAAA)n"
13484_.13531
repeat_region /rpt_family=" (GGGA)n"
complement(14217_.14449)
/rpt_family="Lx"
14656_.14711
repeat_region /rpt_family=" (TC)n"
14711_.14730
repeat_region /rpt_family=" (TG)n"
complement(14880_.15121)
/rpt_family="ORR1D"
15621_.15656
repeat_region /rpt_family=" (CCCCCT)n"
15658_.15841
repeat_region /rpt_family="C-rich"
17085_.17182
repeat_region /rpt_family="PB1D10"
17647_.17694
repeat_region /rpt_family=" (CA)n"
17699_.17905
repeat_region /rpt_family="MLT1C"
complement(18654_.18823)
/rpt_family="RMER10B"
18824_.18902
repeat_region /rpt_family=" (GGAGAA)n"
18903_.19084
repeat_region /rpt_family=" (GGAGAA)n"

18962_.19013
/note="single clone coverage"
19014_.19015
/note="unresolved VNTR: restriction enzyme digest
fingerprint indicates <100 bp of poly [GA] missing here"
complement(19165_.19404)
/rpt_family="RMER10B"
19808_.20229
repeat_region /rpt_family="RMER17C"
20329_.20956
repeat_region /rpt_family="Lx8"
21021_.21066
repeat_region /rpt_family=" (TG)n"

Query Match 76.8%; Score 19.2; DB 9; Length 197585;
Best Local Similarity 87.5%; Pred. No. 2.6e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 AAAAACTGGATCTTCAGGCTGAGA 25
|||||
Db 49411 AAAGACTGGATCAGAGGCTGAGA 49388
|||||

RESULT 49
AC118371/c
LOCUS
DEFINITION
AC118371
AC118371.3 GI:23269853
HTG: HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
Rattus norvegicus (Norway rat)
SOURCE
ORGANISM
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Murinae; Rattus.

REFERENCE
1 (bases 1 to 197966)
Muzny, D., Marie, M., Metsker, M., Lee, A., Abranzon, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,
Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
Davila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
Egan, A., Escotto, M., Evans, C., Evans, C.A., Falls, T., Fan, G.,
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
Gebrgeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hoques, M.,
Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
Lorensheue, L., Loulseged, H., Lozado, R.J., Lu, X., Ma, J.,
Maheshwari, M., Mahindartine, M., Mahmoud, M., Malloy, K., Mangum, A.,
Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
Mawhiney, S., McLeod, M.P., McNeill, T.Z., Meenen, E.,
Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,
Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Parks, K.,
Nwaokemele, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K.,
Pasternak, S., Paul, H., Perez, A., Perez, L., Prannkoch, C.,
Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L.,
Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R.,
Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
```


Insert size: 211557; sum-of-contigs
Quality coverage: 3.94x in Q20 bases; agarose-fp
Quality coverage: 3.96x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 19 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
* as soon as it is available and the accession number will
* be preserved.

* 1 2193: contig of 2193 bp in length
* 2194 2293: gap of unknown length
* 2294 4753: contig of 2460 bp in length
* 4754 4853: gap of unknown length
* 4854 7412: contig of 2559 bp in length
* 7413 7512: gap of unknown length
* 7513 10881: contig of 3369 bp in length
* 10882 10981: gap of unknown length
* 10982 15261: contig of 4280 bp in length
* 15262 15361: gap of unknown length
* 15362 19915: contig of 4554 bp in length
* 19916 20015: gap of unknown length
* 20016 24367: contig of 4352 bp in length
* 24368 24467: gap of unknown length
* 24468 31449: contig of 6982 bp in length
* 31450 31549: gap of unknown length
* 31550 39528: contig of 7979 bp in length
* 39529 39628: gap of unknown length
* 39629 44857: contig of 5228 bp in length
* 44858 44957: gap of unknown length
* 44958 51499: contig of 6343 bp in length
* 51500 51599: gap of unknown length
* 51600 61361: contig of 9762 bp in length
* 61362 61461: gap of unknown length
* 61462 69455: contig of 7993 bp in length
* 69456 69555: gap of unknown length
* 69556 85346: contig of 15792 bp in length
* 85347 85447: gap of unknown length
* 85448 96665: contig of 11219 bp in length
* 96666 96765: gap of unknown length
* 96766 113833: contig of 17068 bp in length
* 113834 132574: gap of unknown length
* 132575 132674: contig of 18641 bp in length
* 132675 132675: gap of unknown length
* 132676 170486: contig of 37812 bp in length
* 170487 170586: gap of unknown length
* 170587 213357: contig of 42771 bp in length.

FEATURES

source

Location/Qualifiers
1. .213357
/organism="Oryctolagus cuniculus"
/mol_type="genomic DNA"
/db_xref="taxon:9986"
/clone_lib="LB1-280H21"
/clone_lib="LB1"
/note="BAC resource: <http://bacpac.chori.org/>
breed: New Zealand White
1. .2193
/note="assembly_fragment"

misc_feature

1. .2193

2194. .2293

gap /estimated_length=unknown

2294. .4753

misc_feature /note="assembly_fragment"

4754. .4853

gap /estimated_length=unknown

4854. .7412

misc_feature /note="assembly_fragment"

7413. .7512

gap /estimated_length=unknown

7513. .10881

misc_feature /note="assembly_fragment"

10882. .10981

gap

misc_feature /estimated_length=unknown
10982. .15261
note="assembly_fragment"
15262. .15361
gap /estimated_length=unknown
15362. .19915
note="assembly_fragment"
19916. .20015
gap /estimated_length=unknown
20016. .24367
note="assembly_fragment"
24368. .24467
gap /estimated_length=unknown
24468. .31449
note="assembly_fragment"
31450. .31549
gap /estimated_length=unknown
31550. .39528
note="assembly_fragment"
39529. .39628
gap /estimated_length=unknown
39629. .44856
note="assembly_fragment"
44857. .44956
gap /estimated_length=unknown
44957. .51499
note="assembly_fragment"
51500. .51599
gap /estimated_length=unknown
51600. .61361
note="assembly_fragment"
61362. .61461
gap /estimated_length=unknown
61462. .69454
note="assembly_fragment"
69455. .69554
gap /estimated_length=unknown
69555. .85346
note="assembly_fragment"
85347. .85446
gap /estimated_length=unknown
85447. .96665
note="assembly_fragment"
96666. .96765
gap /estimated_length=unknown
96766. .113833
note="assembly_fragment"
113834. .113933
gap /estimated_length=unknown
113934. .132574
note="assembly_fragment"
132575. .132674
gap /estimated_length=unknown
132675. .170486
note="assembly_fragment"
170487. .170586
gap /estimated_length=unknown
170587. .213357
note="assembly_fragment"

ORIGIN

Query Match 76.8%; Score 19.2; DB 14; Length 213357;
Best Local Similarity 87.5%; Pred No. 2.5e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAACTGGAATCTCAGGCTGAG 24
||||| ||||| ||||| |||||

Db 195552 AAAACCTGCAATCCAGGCTGAG 195529
||||| ||||| ||||| |||||

Search completed: February 3, 2006, 21:27:26
Job time : 876 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 3, 2006, 10:53:33 ; Search time 203.111 Seconds
(without alignments)
820.326 Million cell updates/sec

Title: US-10-719-900-30

Perfect score: 25

Sequence: 1 aaaaactggaatctcaggtgaga 25

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 333246308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database :

N Geneseq_21.*
1: Geneseqn1980s.*
2: Geneseqn1980s.*
3: Geneseqn2000s.*
4: Geneseqn2001bs.*
5: Geneseqn2001bs.*
6: Geneseqn2002bs.*
7: Geneseqn2002bs.*
8: Geneseqn2003as.*
9: Geneseqn2003bs.*
10: Geneseqn2003cs.*
11: Geneseqn2003ds.*
12: Geneseqn2004as.*
13: Geneseqn2004bs.*
14: Geneseqn2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	80.0	300	2	Aaz12748 Human gen
2	20	80.0	758	6	Aaz15676 Human gen
3	20	80.0	5174	6	Abn59811 Novel hum
4	20	80.0	11124	10	Adc19453 cDNA encod
5	20	80.0	11127	14	Ady17371 DNA encod
6	18.6	74.4	1484	2	Aax20314 Borrelia
7	18.6	74.4	2065	10	Adas2903 Human cod
8	18.6	74.4	2167	10	Add45029 Rat gene
9	18	72.0	956	8	Abz18441 Group III
10	17.8	71.2	876	13	Adx30549 Plant ful
11	17.8	71.2	1911	13	Adsa4564 Bacterial
12	17.8	71.2	28170	6	Abao1447 Streptoco
13	17.6	70.4	214	3	Adf57295 Urogenita
14	17.6	70.4	276	12	Adp28801 Human sec
15	17.6	70.4	522	12	Ach68960 Human gen
16	17.6	70.4	561	4	Aah09027 Human cdn
17	17.6	70.4	1266	13	Acn39840 Tumour-as
18	17.6	70.4	1688	4	Aah13739 Human cdn
19	17.6	70.4	2177	13	Acn39839 Tumour-as

c 20	70.4	3081	13	ADP07065	Adp07065 Full leng
c 21	70.4	5271	8	ABX63354	ABX63354 Human cdn
c 22	70.4	5271	10	ADJ56534	ADJ56534 Rat cdna
c 23	70.4	14815	6	ABS76688	ABS76688 PRRS viru
c 24	70.4	20555	14	ADZ113631_4	Continuation (5 of
c 25	70.4	20555	14	ADZ13620_4	Continuation (5 of
c 26	70.4	24526	6	ADA43979	Ada43979 Human tra
c 27	70.4	24526	13	ADU48434	Adu48434 Human tra
c 28	70.4	26320	10	ABZ24431	Abz24431 Human con
c 29	70.4	37442	10	ADZ59528	Adz59528 Secondary
c 30	70.4	37442	14	ADZ59517	Adz59517 Secondary
c 31	70.4	37442	14	ABE96540	ABE96540 Human ill
c 32	70.4	37996	4	ABL07876	AbL07876 Drosophill
c 33	70.4	49999	2	AAZ23896	Aaz23896 Murine LO
c 34	70.4	49999	2	AAZ23891	Aaz23891 Murine LO
c 35	70.4	63411	12	ADQ97081	Adq97081 Mouse can
c 36	70.4	96594	10	ADC85257	Adc85257 Mouse ptp
c 37	70.4	96595	9	ADA02777	Ada02777 Mouse ptp
c 38	70.4	96595	10	ADB72515	AdB72515 Mouse ptp
c 39	70.4	96595	12	ADM74372	Adm74372 Murine ca
c 40	70.4	110000	11	ACN43998_2	Continuation (3 of
c 41	70.4	110000	11	ACN45090_0	Continuation (3 of
c 42	70.4	154902	6	ABQ88198	Abq88198 Human oet
c 43	70.4	289106	14	ADZ12614	Adz12614 Murine ca
c 44	70.4	299598	12	ADQ59380	Adq59380 Human can
c 45	70.4	1059	13	ADO81499	Ado81499 Plant ful
c 46	70.4	2206	13	ADO83966	Ado83966 Plant ful
c 47	70.4	2800	13	ADX63913	Adx63913 Plant ful
c 48	70.4	5161	5	ABA20105	AbA20105 Human ner
c 49	70.4	35456	11	ACN43916	Acn43916 Mouse gen
c 50	70.4	460	12	ADP94079	Adp94079 Cotton ex
c 51	70.4	517	10	ADB49718	AdB49718 Primary r
c 52	70.4	517	13	ADV39311	Adv39311 Rat card
c 53	70.4	877	2	AAZ16499	Aaz16499 Human gen
c 54	70.4	944	13	ADR64470	Adr64470 Cotton cd
c 55	70.4	1448	13	ADR64471	Adr64471 Cotton cd
c 56	70.4	1935	4	AAK67546	Aak67546 Human imm
c 57	70.4	2372	5	ADL63428	AdL63428 Human ova
c 58	70.4	3199	13	ADR07384	Adr07384 Full leng
c 59	70.4	3670	13	ADR08219	Adr08219 Full leng
c 60	70.4	4047	12	ADN00348	Adn00348 Novel hum
c 61	70.4	4683	8	ABT23216	Abt23216 Human pro
c 62	70.4	4683	13	ADR06893	Adr06893 Full leng
c 63	70.4	5054	12	ADQ22381	Adq22381 Human sof
c 64	70.4	6901	13	ADR84230	Adr84230 Aspergill
c 65	70.4	33486	6	ABS78976	AbS78976 E. coli C
c 66	70.4	33486	10	ADH80543	Adh80543 Escherich
c 67	70.4	35236	11	ACN44094	Acn44094 Human gen
c 68	70.4	47841	6	ABQ80553	Abq80553 Human can
c 69	70.4	76180	13	ABD33385	ABd33385 Human can
c 70	70.4	80959	8	AAZ151405	Aaz151405 Human sec
c 71	70.4	143239	12	ADQ17729	Adq17729 Human sof
c 72	70.4	222930	6	ABR84349	AbR84349 Human cdn
c 73	70.4	136	4	AAZ50021	Aaz50021 Staphyloc
c 74	70.4	136	8	ACA17270	AcA17270 Prokaryot
c 75	70.4	193	4	AAI25608	Aai25608 Probe #15
c 76	70.4	193	4	ABA71858	AbA71858 Human foe
c 77	70.4	193	4	AAI52202	Aai52202 Probe #20
c 78	70.4	193	4	ABA37906	AbA37906 Probe #16
c 79	70.4	193	4	AAK46303	Aak46303 Human bron
c 80	70.4	193	4	AAK20234	Aak20234 Human bra
c 81	70.4	193	4	ABS46026	AbS46026 Human liv
c 82	70.4	193	6	ABS20617	AbS20617 Human gen
c 83	70.4	300	2	AAZ14445	Aaz14445 Human gen
c 84	70.4	304	3	ABQ19979	AbQ19979 Human sec
c 85	70.4	349	6	ABQ85851	Abq85851 Arabidops
c 86	70.4	388	6	ABN60231	Abn60231 Human can
c 87	70.4	397	4	AAI81444	Aai81444 Human pol
c 88	70.4	399	4	AAAS48855	Aas48855 Staphyloc
c 89	70.4	399	8	ACA16126	AcA16126 Prokaryot
c 90	70.4	402	8	ABX43585	Abx43585 Bovine ES
c 91	70.4	416	6	ABN24330	Abn24330 Human ORF
c 92	70.4	456	4	AAI16428	Aai16428 Probe #63

XX 12-OCT-1999 (first entry)
DT Human gene expression product cDNA sequence SEQ ID NO:3145.
XX
XX Human; gene; gene expression product; diagnosis; therapy; probe;
KW detection; mapping; tissue typing; profiling; forensic; cancer;
KW genetic analysis; colorectal cancer; breast cancer; lung cancer; ss.
XX
XX Homo sapiens.
OS
XX WO9938972-A2.
PN
XX 05-AUG-1999.
PD
XX
XX 28-JAN-1999; 99WO-US001619.
PF
XX
XX 28-JAN-1998; 98US-0072910P.
PR
XX 24-FEB-1998; 98US-0075954P.
PR
XX 31-MAR-1998; 98US-0080114P.
PR
XX 03-APR-1998; 98US-0080515P.
PR
XX 01-APR-1998; 98US-0080666P.
PR
XX 21-OCT-1998; 98US-0105234P.
PR
XX 28-OCT-1998; 98US-0105877P.
PR
XX
XX (CHIR) CHIRON CORP.
PA (HYSE-) HYSEQ INC.
PA
XX Williams LT, Escobedo J, Innis MA, Garcia PD, Sudduth-Klinger J;
PI Reinhard C, Giese K, Randazzo F, Kennedy GC, Pot D, Kassam A;
PI Lamson G, Drmanac R, Crkvenjakov R, Dickson M, Drmanac S, Labat I;
PI Leshkowitz D, Kita D, Garcia V, Jones WL, Stache-Crain B;
XX
XX WPI; 1999-494092/41.
DR
XX
XX Novel human genes and their expression products which are differentially
PT expressed in different cell types.
PT
XX
XX Claim 1; Page 1511; 2479pp; English.
PS
XX
XX The present invention describes a library of human polynucleotides
CC comprising the sequences given in AA212532 to AA217779. Also described is
CC a method of detecting differentially expressed genes correlated with the
CC cancerous state of a mammalian cell, comprising detecting at least one
CC differentially expressed gene product in a test sample from a cell
CC suspected of being cancerous, where the gene product is encoded by one of
CC the 5248 polynucleotide sequences given in AA212532 to AA217779. The
CC polynucleotides can be used as a source of primers and probes, which can
CC be used for a variety of purpose, e.g. detection of expression levels,
CC mapping, tissue typing or profiling, forensics, genetic analysis and
CC detection of polymorphisms. Polypeptides encoded by the polynucleotides
CC can be used for raising antibodies for experimental, diagnostic and
CC therapeutic purposes. The polynucleotides may also be used to construct
CC arrays for diagnostics (which may be used to determine function of an
CC encoded protein); and to detect differences in expression levels between
CC two cells (e.g. to identify abnormal or diseased tissue in a human, to
CC identify a genetic predisposition or susceptibility to a disease such as
CC cancer). The polynucleotides of the invention are especially used in the
CC diagnosis, prognosis and management of colorectal cancer, breast cancer,
CC and lung cancer. The polynucleotides can also be used to screen for
CC peptide analogues and antagonists
XX
SQ Sequence 758 BP; 221 A; 146 C; 148 G; 215 T; 0 U; 28 Other;
Query Match 80.0%; Score 20; DB 2; Length 758;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 4 AAACCTGGAATCTCAGGCTGA 23
Db 148 AAACCTGGAATCTCAGGCTGA 167

RESULT 3
ABN59811
ID ABN59811 standard; cDNA; 5174 BP.
XX
XX AC ABN59811;
XX
XX 28-JUN-2002 (first entry)
DT
XX
XX Novel human coding sequence SEQ ID NO: 222.
DE
XX
XX Human; antianaemic; vulnery; antiinflammatory; immunomodulator;
KW antifertility; cerebroprotective; cytostatic; rheumatic; gene therapy;
KW neuroprotective; antiparkinsonian; protein therapy; EST;
KW expressed sequence tag; gene; ss.
XX
XX Homo sapiens.
OS
XX WO200222660-A2.
PN
XX 21-MAR-2002.
PD
XX 10-SEP-2001; 2001WO-US026015.
PF
XX 11-SEP-2000; 2000US-00659671.
PR
XX (HYSE-) HYSEQ INC.
PA
XX Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
PI Xue AJ, Yang Y, Wehrman T, Drmanac RT;
PI
XX WPI; 2002-292408/33.
DR
XX P-PSDB; ABB97398.
DR
XX
XX An isolated polynucleotide for treating diseases associated with its
PT encoded polypeptide such as cancer and multiple sclerosis.
PT
XX
XX Claim 1; SEQ ID NO 222; 509pp; English.
PS
XX
XX The present invention provides the protein and coding sequences of 444
CC novel human proteins. These were isolated from expressed sequences tags
CC (ESTs). They can be used to stimulate cell growth, to regulate
CC haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth
CC e.g. in burn treatment, to regulate the immune system e.g. to treat
CC multiple sclerosis, to regulate activin or inhibin e.g. to treat
CC infertility, to regulate haemostasis or thrombolysis e.g. to treat stroke
CC and cancer, to screen for drugs, to treat inflammatory conditions e.g.
CC rheumatoid arthritis, and to treat nervous system disorders e.g. of
CC Parkinson's disease. The present sequence is a coding sequence of the
CC invention
XX
SQ Sequence 5174 BP; 1361 A; 1184 C; 1262 G; 1367 T; 0 U; 0 Other;
Query Match 80.0%; Score 20; DB 6; Length 5174;
Best Local Similarity 100.0%; Pred. No. 53;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 4 AAACCTGGAATCTCAGGCTGA 23
Db 964 AAACCTGGAATCTCAGGCTGA 983

RESULT 4
ACD19453
ID ACD19453 standard; cDNA; 11124 BP.
XX
XX AC ACD19453;
XX
XX 25-AUG-2003 (first entry)
DT
XX
XX cDNA encoding novel human protein #133.
DE
XX
XX Human; NOV; gene therapy; endocrine related disease; diabetes;
KW metabolism-related disease; obesity; central nervous system disorder;

KW Alzheimer's disease; Parkinson's disease; epilepsy; multiple sclerosis;
KW schizophrenia; depression; autoimmune disorder; inflammatory disorder;
KW peoriasis; allergy; lupus erythematosus; asthma; cancer;
KW inflammatory bowel disease; rheumatoid arthritis; osteoarthritis;
KW colon cancer; lung cancer; liver cancer; breast cancer; ovarian cancer;
KW prostate cancer; brain cancer; melanoma; liver disease; liver cirrhosis;
KW lung disease; emphysema; obstructive pulmonary disease; haemophilia;
KW stroke; infection; gene; ss.

OS Homo sapiens.

XX
XX WO2003023002-A2.

XX
XX 20-MAR-2003.

XX
XX 09-SEP-2002; 2002WO-US028539.

XX
XX 07-SEP-2001; 2001US-0318120P.

XX
XX 07-SEP-2001; 2001US-0318130P.

XX
XX 10-SEP-2001; 2001US-0318430P.

XX
XX 17-SEP-2001; 2001US-0322836P.

XX
XX 17-SEP-2001; 2001US-0322781P.

XX
XX 17-SEP-2001; 2001US-0322816P.

XX
XX 17-SEP-2001; 2001US-0322817P.

XX
XX 19-SEP-2001; 2001US-0323519P.

XX
XX 20-SEP-2001; 2001US-0323631P.

XX
XX 20-SEP-2001; 2001US-0323636P.

XX
XX 25-SEP-2001; 2001US-0324969P.

XX
XX 25-SEP-2001; 2001US-0325091P.

XX
XX 26-SEP-2001; 2001US-0324990P.

XX
XX 17-APR-2002; 2002US-0373212P.

XX
XX 06-SEP-2002; 2002US-00236177.

XX
XX (CURA-) CURAGEN CORP.

XX
XX Spytek KA, Patturajan M, Gorman L, Li L, Anderson DW, Zhong M;

XX Gerlach VL, Vernet CAM, Ellerman K, Berghs C, Rothenberg ME, Guo X;

XX Shimkets RA, Leach MD, Catterton E, Kekuda R, Ji W, Miller CE;

XX Riegler DK, Taupier RJ, Shenoy SG, Liu X, Padigaru M, Alsobrook JP;

XX Lepley DM, Edinger SR, Burgess CE;

XX
XX WPI; 2003-313242/30.

XX
XX P-PSDB; ABO14760.

XX New cytoplasmic, nuclear membrane bound or secreted polypeptides (NOVX)
XX and polynucleotides, useful in gene therapy, e.g. for treating or
XX preventing obesity, multiple sclerosis, allergy, cancers, hemophilia,
XX stroke or infections.

XX
XX Claim 20; Page 373-375; 586pp; English.

XX The invention describes a new isolated polypeptide (NOVX). The NOVX
XX polypeptide, nucleic acid and antibody are useful as therapeutics,
XX particularly in the manufacture of a medicament for treating a syndrome
XX associated with a human disease, which includes a pathology associated
XX with NOVX polypeptide. The DNA encoding the protein is useful in gene
XX therapy for treating the disease or condition. In particular, the NOVX
XX polypeptide or polynucleotide is useful for treating endocrine/
XX metabolism-related diseases (e.g. obesity or diabetes), central nervous
XX system disorders (e.g. Alzheimer's disease, Parkinson's disease,
XX epilepsy, multiple sclerosis, schizophrenia or depression), autoimmune
XX and inflammatory disorders (e.g. psoriasis, allergy, lupus erythematosus,
XX asthma, inflammatory bowel disease, rheumatoid arthritis or
XX osteoarthritis), cancers (e.g. colon, lung, liver, breast, ovarian,
XX prostate or brain cancers, or melanoma), liver diseases (e.g. liver
XX cirrhosis), lung diseases (emphysema or obstructive pulmonary disease),
XX haemophilia, stroke, or infections (e.g. viral, bacterial or parasitic).
XX These are also useful in developing powerful assay system for functional
XX analysis of various human disorders, as well as in diagnostic
XX applications, and for monitoring the effects of drugs during clinical
XX trials. This sequence encodes a novel human NOV protein

XX Sequence 11124 BP; 2855 A; 2715 C; 2624 G; 2928 T; 0 U; 2 Other;

Query Match 80.0%; Score 20; DB 10; Length 11124;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 AAACCTGGAATCTCAGGCTGA 23

|||||

Db 964 AAACCTGGAATCTCAGGCTGA 983

RESULT 5

ADY17371

ID ADY17371 standard; DNA; 11127 BP.

XX

AC ADY17371;

XX
XX 05-MAY-2005 (first entry)

XX
XX DNA encoding a PRO polypeptide, SEQ ID NO 3177.

XX Antinflammatory; Immune disorder; Dermatological; Immunosuppressive;
XX Antirheumatic; Antiarthritic; Osteopathic; Hemostatic; Antianemic;
XX Antithyroid; Antidiabetic; Nephrotropic; CNS-Gen.; Hepatotropic;
XX Virucide; Gastrointestinal-Gen.; Antipsoriatic; Antiasthmatic;
XX Antiallergic; ds; gene; diagnosis.

XX
XX Homo sapiens.

XX
XX WO2005016962-A2.

XX
XX 24-FEB-2005.

XX
XX 11-AUG-2004; 2004WO-US026249.

XX
XX 11-AUG-2003; 2003US-0493546P.

XX
XX (GETH) GENENTECH INC.

XX
XX Abbas A, Clark H, Ouyang W, Williams MP, Wood WI, Wu TD;

XX
XX WPI; 2005-182330/19.

XX New nucleic acid encoding PRO polypeptide, useful for diagnosing and
XX treating an immune related disorder, e.g. systemic lupus erythematosus,
XX rheumatoid arthritis, osteoarthritis, thyroiditis, or diabetes mellitus.

XX
XX Claim 1; SEQ ID NO 3177; 158pp; English.

XX The invention relates to an isolated nucleic acid encoding a PRO
XX polypeptide. The polypeptide, agonist or an antagonist, antibody,
XX composition, and method are useful for diagnosing and treating an immune
XX related disorder, e.g. systemic lupus erythematosus, rheumatoid
XX arthritis. The present sequence represents a DNA encoding a PRO
XX polypeptide.

XX
XX Sequence 11127 BP; 2858 A; 2712 C; 2628 G; 2929 T; 0 U; 0 Other;

Query Match 80.0%; Score 20; DB 14; Length 11127;

Best Local Similarity 100.0%; Pred. No. 60;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 AAACCTGGAATCTCAGGCTGA 23

|||||

Db 964 AAACCTGGAATCTCAGGCTGA 983

RESULT 6

AA20314/C

ID AAX20314 standard; DNA; 1484 BP.

XX
XX AAX20314;

XX
XX 04-MAY-1999 (first entry)

DT

```
XX Borrelia burgdorferi polynucleotide sequence #67.
DE
XX
XX Borrelia burgdorferi; spirochete; bacterium; pathogen; Lyme disease;
KW epidemic relapsing fever; endemic relapsing fever; Lyme borreliosis;
KW infection; diagnosis; characterisation; detection; ds.
XX
XX Borrelia burgdorferi.
OS
XX
XX WO9858943-A1.
PN
XX
XX 30-DEC-1998.
PD
XX
XX 18-JUN-1998; 98WO-US012764.
PF
XX
XX 20-JUN-1997; 97US-0050359P.
XX
XX 22-JUL-1997; 97US-0053344P.
PR
XX
XX 22-JUL-1997; 97US-0053377P.
PR
XX
XX 03-SEP-1997; 97US-0057483P.
PR
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX (MEDI-) MEDIMUNE INC.
PA
XX
XX Fraser C, White OR, Clayton R, Dougherty BA, Lathigra R;
PI Smith HO;
XX
XX WPI; 1999-081217/07.
XX
XX New isolated Borrelia burgdorferi nucleic acids - used to develop
PT products for the detection, diagnosis, characterisation, prevention and
PT therapy of infections, particularly Lyme disease.
XX
XX Claim 1; Page 1050-1051; 1128pp; English.
XX
XX AAX20248 to AAX20402 represent polynucleotide sequences isolated from
CC Borrelia burgdorferi (Bb). Products derived from Bb can be used for the
CC detection, diagnosis, characterisation, prevention and therapy of Bb
CC infections, e.g. Lyme disease. They can also be used for the production
CC of biosynthetic products, e.g. enzymes. Borrelia belongs to a family of
CC motile, spiral-shaped bacteria called Spirochetes. Spirochetes are
CC pathogenic in humans and Borrelia causes epidemic and endemic relapsing
CC fever, and Lyme borreliosis, more commonly known as Lyme disease
XX
XX Sequence 1484 BP; 410 A; 290 C; 208 G; 571 T; 0 U; 5 Other;
SQ
Query Match 74.4%; Score 18.6; DB 2; Length 1484;
Best Local Similarity 84.0%; Pred. No. 1.9e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

XX
XX 1 AAAAACTGGAATCTCAGGCTGAGA 25
Db 454 AAAAACTGCAAGAACAGGCTGAGA 430

RESULT 7
ADA52903
ID ADA52903 standard; cDNA; 2066 BP.
XX
XX ADA52903;
AC
XX
XX 20-NOV-2003 (first entry)
DT
XX
XX Human coding sequence, SEQ ID 471.
DE
XX
XX Cytostatic; Anti-inflammatory; Osteopathic; Neuroprotective; Nootropic;
KW Gene Therapy; human; secretory protein; membrane proteins; cancer;
KW inflammatory disease; osteoporosis; neurological disease; gene; ss.
XX
XX Homo sapiens.
OS
XX
XX EP1293569-A2.
PN
XX
XX 19-MAR-2003.
PD
```

```
XX 21-MAR-2002; 2002EP-00006586.
PF
XX
XX 14-SEP-2001; 2001JP-00328381.
PR
XX
XX 24-JAN-2002; 2002US-0350435P.
PR
XX
XX (HELI-) HELIX RES INST.
PA
XX
XX (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
XX Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tameshika I;
PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;
XX
XX WPI; 2003-395539/38.
DR
XX
XX P-PSDB; ADA54542.
DR
XX
XX New polynucleotides encoding full-length polypeptides, e.g. secretory
PT and/or membrane proteins, useful for developing medicines for diseases in
PT which the gene is involved, or as target molecules for gene therapy.
XX
XX Claim 1; SEQ ID NO 471; 205pp; English.
XX
XX The present invention relates to novel human secretory or membrane
CC proteins (ADA54072-ADA55710) and their coding sequences (ADA52433-
CC ADA54071). The coding sequences are useful in the gene therapy of
CC diseases caused by abnormalities of the proteins, e.g. cancer,
CC inflammatory diseases, osteoporosis or neurological disease.
XX
XX Sequence 2066 BP; 460 A; 696 C; 537 G; 373 T; 0 U; 0 Other;
SQ
Query Match 74.4%; Score 18.6; DB 10; Length 2066;
Best Local Similarity 84.0%; Pred. No. 2e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

XX
XX 1 AAAAACTGGAATCTCAGGCTGAGA 25
Db 820 AAAAACTCGAAACTCGGCTGAGA 844

RESULT 8
ADA5029/c
ID ADA5029 standard; DNA; 2167 BP.
XX
XX ADA5029;
AC
XX
XX 29-JAN-2004 (first entry)
DT
XX
XX Rat gene Z48444, SEQ ID NO 10461.
DE
XX
XX Rat; ds; gene; pain; neuronal tissue; gene therapy;
KW spinal segmental nerve injury; chronic constriction injury; CCI;
KW spared nerve injury; SNI; Chung.
XX
XX Rattus norvegicus.
OS
XX
XX WO2003016475-A2.
PN
XX
XX 27-FEB-2003.
PD
XX
XX 14-AUG-2002; 2002WO-US025765.
XX
XX 14-AUG-2001; 2001US-0312147P.
PR
XX
XX 01-NOV-2001; 2001US-0346382P.
PR
XX
XX 26-NOV-2001; 2001US-033347P.
PR
XX
XX (GEHO) GEN HOSPITAL CORP.
XX
XX (FARB) BAYER AG.
PA
XX
XX Woolf C, D'urso D, Befort K, Costigan M;
PI
XX
XX WPI; 2003-268312/26.
DR
XX
XX GENBANK; Z48444.
XX
```

PT New composition comprising two or more isolated polypeptides, useful for
 PT preparing a medicament for treating pain in an animal.
 XX
 PS Claim 1; Page; 1017pp; English.

XX The invention discloses a composition comprising two or more isolated rat
 CC or human polynucleotides or a polynucleotide which represents a fragment,
 CC derivative or allelic variation of the nucleic acid sequence. Also
 CC claimed are a vector comprising the novel polynucleotide, a host cell
 CC comprising the vector, a method for identifying a nucleotide sequence
 CC which is differentially regulated in an animal subjected to pain and a
 CC kit to perform the method, an array, a method for identifying an agent
 CC that increases or decreases the expression of the polynucleotide sequence
 CC that is differentially expressed in neuronal tissue of a first animal
 CC subjected to pain, a method for identifying a compound which regulates
 CC the expression of a polynucleotide sequence which is differentially
 CC expressed in an animal subjected to pain, a method for identifying a
 CC compound that regulates the activity of one or more of the
 CC polynucleotides, a method for producing a pharmaceutical composition, a
 CC method for identifying a compound or small molecule that regulates the
 CC activity in an animal of one or more of the polypeptides given in the
 CC specification, a method for identifying a compound useful in treating
 CC pain and a pharmaceutical composition comprising the one or more
 CC polypeptides or their antibodies. The polynucleotide or the compound that
 CC modulates its activity is useful for preparing a medicament for treating
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
 CC therapy). The sequence presented is a rat DNA (shown in Table 2 of the
 CC specification) which encodes one of the polypeptides of the invention
 CC which is differentially expressed during pain. Note: The sequence data
 CC for this patent did not form part of the printed specification, but was
 CC obtained in electronic form directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX
 SQ Sequence 2167 BP; 611 A; 441 C; 510 G; 605 T; 0 U; 0 Other;

Query Match 74.4%; Score 18.6; DB 10; Length 2167;
 Best Local Similarity 84.0%; Pred.No. 2.1e+02;
 Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAACTGGATCTCAGGCTGAGA 25
 ||||| ||||| ||||| ||||| |||||
 DB 1642 AAAAATTGGAATTCAGGCTTAAA 1618

RESULT 9
 ABZ18441
 ID ABZ18441 standard; cDNA; 956 BP.

AC ABZ18441;

XX 23-JAN-2003 (first entry)

XX Group III CDNA cancer related clone SEQ ID NO:867.

XX Human; cancer; tumour; therapy; diagnosis; CT antigen; CP antigen;
 KW immune response; virology; immunology; microbiology; molecular biology;
 KW recombinant DNA technology; gene; ss.

XX Homo sapiens.

XX WO200278516-A2.

XX 10-OCT-2002.

XX 28-MAR-2002; 2002NO-US010421.

XX 30-MAR-2001; 2001US-0280255P.

PR 28-AUG-2001; 2001US-0315563P.

PR 09-JAN-2002; 2002US-0347313P.

PA (CORI-) CORIXA CORP.

PI Wang T, Wang S, Bangur CS, Gaiger A;
 XX WPI; 2003-058387/05.

XX New immunogenic polynucleotides or polypeptides useful for diagnosing,
 PT preventing and treating cancer expressing CT or CP mRNA antigens, and in
 PT virology, immunology, microbiology, molecular biology and recombinant DNA
 PT techniques.

PS Claim 1; SEQ ID NO 867; 207pp; English.

XX ABQ17575 to ABQ20506 represent isolated polynucleotide (I) sequences, and
 CC ABP54446 to ABP54472 represent protein (II) sequences, from the present
 CC invention. (I) and (II) have cytostatic activity and can be used in gene
 CC therapy and vaccines (I), (II), antibodies and compositions from the
 CC present invention are useful for diagnosing, preventing and treating
 CC cancer, which expresses CT or CP mRNA antigens. They are useful for
 CC stimulating immune response. They can also be useful in virology,
 CC immunology, microbiology, molecular biology and recombinant DNA
 CC techniques. N.B. The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 956 BP; 273 A; 168 C; 189 G; 301 T; 0 U; 25 Other;

Query Match 72.0%; Score 18; DB 8; Length 956;
 Best Local Similarity 100.0%; Pred.No. 3.4e+02;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 AAAACTGGAATCTCAGGC 20
 ||||| ||||| ||||| ||||| |||||
 DB 687 AAAACTGGAATCTCAGGC 704

RESULT 10
 ADX30549/c
 ID ADX30549 standard; cDNA; 876 BP.

XX AC ADX30549;

XX 21-APR-2005 (first entry)

XX Plant full length insert polynucleotide seqid 13369.

XX plant protectant; plant growth regulant; gene therapy; plant;
 KW recombinant DNA construct; physical array; plant breeding marker;
 KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
 KW extreme osmotic condition; pathogen tolerance; pest tolerance;
 KW growth rate; cell cycle pathway; disease resistance;
 KW galactomannan production; lignin production; plant growth regulator;
 KW yield; plant growth; plant development; seed oil; protein yield;
 KW protein content; gene; ss.

XX Unidentified.

XX US2004034888-A1.

XX 19-FEB-2004.

XX 28-APR-2003; 2003US-00425114.

XX 06-MAY-1999; 99US-00304517.

PR 05-NOV-2001; 2001US-00985678.

XX (LIUJ/) LIU J.

PA (ZHOU/) ZHOU Y.

PA (KOVA/) KOVALIC D K.

PA (SCRE/) SCREEN S B.

PA (TABA/) TABASKA J E.

XX (CAOY/) CAO Y.

XX Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;

DR WPI; 2004-180133/17.
XX
XX New recombinant DNA construct, useful for improving plant tolerance to
PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
PT pests, for conferring increased resistance to plant disease, or for
PT improving yield.
XX
XX Claim 1; SEQ ID NO 13369; 15pp; English.
XX
XX The invention describes a recombinant DNA construct comprising a
CC polynucleotide consisting of a sequence encoding an amino acid sequence
CC available in electronic form from the US patent office at
CC ftp.segdata.uspto.gov/sequence.html?DocID:2004034888. The polynucleotide
CC of the invention are also useful in physical arrays of molecules and as
CC plant breeding markers. The recombinant DNA construct is useful for
CC improving plant tolerance to cold, heat, drought, herbicides, extreme
CC osmotic conditions, pathogens or pests, for manipulating growth rate in
CC plant cells by modification of the cell cycle pathway, for conferring
CC increased resistance to plant disease, for producing galactomannan,
CC lignin or plant growth regulators, for increasing the rate of homologous
CC recombination in plants, for improving yield by modification of
CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
CC or by providing improved plant growth and development under at least one
CC stress condition or for modifying seed oil or protein yield and/or
CC content. This sequence represents a plant full length insert
CC polynucleotide that can be used in the recombinant DNA construct of the
CC invention.
XX
SQ Sequence 876 BP; 257 A; 145 C; 193 G; 281 T; 0 U; 0 Other;
Query Match 71.2%; Score 17.8; DB 13; Length 876;
Best Local Similarity 90.5%; Pred. No. 4.2e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 4 AACTGGAATCTCAGGCTGAG 24
Db 258 AACTGCAGTCTCAGGCTGAG 238
||||| |||||||
RESULT 11
ADS45644
ID ADS45644 standard; cDNA; 1911 BP.
XX
AC ADS45644;
XX
XX 02-DEC-2004 (first entry)
XX Bacterial polynucleotide #387.
XX Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polynucleotide; gene; ss.
XX Bacteria.
XX
XX US2003233675-A1.
XX
XX 18-DEC-2003.
XX
XX 20-FEB-2003; 2003US-00369493.
XX
XX 21-FEB-2002; 2002US-0360039P.
XX
XX (CAOY/) CAO Y.
PA (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.
XX

PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX WPI; 2004-061375/06.
XX
XX New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX
XX Claim 1; SEQ ID NO 24074; 122pp; English.
XX
XX The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polynucleotide used in
CC the scope of the invention. Note: The sequence data for this patent did
CC not form part of the printed specification but was obtained in electronic
CC format from USPTO at segdata.uspto.gov/sequence.html.
XX
SQ Sequence 1911 BP; 575 A; 431 C; 484 G; 421 T; 0 U; 0 Other;
Query Match 71.2%; Score 17.8; DB 13; Length 1911;
Best Local Similarity 90.5%; Pred. No. 4.8e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 3 AAAACTGGAATCTCAGGCTGA 23
Db 909 AAAACTGGAATCTCGTTGA 929
||||| |||||||
RESULT 12
ABA01447
ID ABA01447 standard; DNA; 28170 BP.
XX
XX ABA01447;
XX
XX 21-FEB-2002 (first entry)
XX Streptococcus thermophilus eps10 operon #2.
XX Exopolysaccharide; lactic acid bacterium; eps, fermented food product;
KW ds.
XX Streptococcus thermophilus.
XX
XX WO200179500-A2.
XX
XX 25-OCT-2001.
XX
XX 18-APR-2001; 2001WO-FR001199.
XX
XX 18-APR-2000; 2000FR-00004972.
XX
XX (INRG) INRA INST NAT RECH AGRONOMIQUE.
PA (DANO-) CIE DANONE SA GERVAIS.
PA (RHOD) RHODIA CHIM.
XX

PI Rallu F, Besancon-Yoshpe I, Fremaux C, Mengaud J, Renault P;
XX WPI; 2002-017616/02.
XX
XX New nucleic acid fragments containing exopolysaccharide operon, useful
PT e.g. for increasing exopolysaccharide synthesis in lactic acid bacteria.
PT
XX
XX Claim 9; Page 109-117; 144pp; French.
XX
XX The present sequence is an eps operon from *Streptococcus thermophilus*.
CC Proteins encoded by the eps operon function in exopolysaccharide (EPS)
CC synthesis. The operon is useful for producing chimeric eps operons, for
CC optimising production of EPS in lactic acid bacteria. EPS impart texture,
CC mouth feel and rheological properties to fermented food products (e.g.
CC yoghurt). They function as thickeners, to provide free-flowing and creamy
CC texture, and may also have biological activities beneficial to health
XX
SQ Sequence 28170 BP; 8951 A; 4464 C; 5468 G; 9287 T; 0 U; 0 Other;
Query Match 71.2%; Score 17.8; DB 6; Length 28170;
Best Local Similarity 90.5%; Pred. No. 7.3e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AAAAAGCTGGAATCTCAGGCTG 22
||||| ||||| ||||| ||||| |||||
Db 7501 AAAAAGCTGGAATCTCAGGCTG 7521

RESULT 13
ADF57295/c
ID ADF57295 standard; cDNA; 214 BP.
XX
XX ADF57295;
XX
DT 12-FEB-2004 (first entry)
XX
DE Urogenital sinus-derived expressed sequence tag, SEQ ID NO 614.
XX
XX foetal urogenital sinus-derived expressed sequence tag; EST; g092; ug093;
KW ug101; ug102; ug106; ug120; ug254; ug291; ug307; ug308; ug311; ug317;
KW ug320; ug334; ug335; ug353; ug354; ug357; ug440; ug441; ug482; ug484;
KW ug485; ug491; ug493; ug494; ug503; ug505; ug506; ug5148; ug5186; ug5194;
KW cytostatic; anti-inflammatory; prostate disease; prostatitis; benign;
KW malignant; prostate gland; ss; mouse; murine.
XX
XX Mus sp.
XX
XX WO9558665-A2.
XX
XX 18-NOV-1999.
XX
XX 14-MAY-1999; 99WO-US010746.
XX
XX 14-MAY-1998; 98US-0085383P.
XX
XX (UYVI-) UNIV VIRGINIA PATENT FOUND.
XX
XX Sikes RA, Chung LWK, Kim JH, Fasciana C, Trapman J;
PI WPI; 2000-039105/03.
XX
XX New urogenital sinus-derived expressed sequence tags, used to, e.g. treat
PT prostate cancer.
PT
XX
XX Disclosure; SEQ ID NO 614; 605pp; English.
XX
XX The invention relates to novel foetal urogenital sinus-derived expressed
CC sequence tags. The invention more specifically comprises an isolated
CC polynucleotide comprising a nucleotide sequence containing a urogenital
CC sinus-derived expressed sequence tag (EST) comprising g092, ug093, ug101,
CC ug102, ug106, ug120, ug254, ug291, ug307, ug308, ug311, ug317, ug320,
CC ug334, ug335, ug353, ug354, ug357, ug440, ug441, ug482, ug484, ug485,
CC ug491, ug493, ug494, ug503, ug505, ug506, ug5148, ug5186 and ug5194. The

CC urogenital sinus-derived expressed sequence tags have cytostatic and anti
CC -inflammatory activities. The genes containing urogenital sinus-derived
CC ESTs encode oncofetal proteins that are useful for diagnosis of prostate
CC disease and for monitoring the efficacy of a compound in clinical trials
CC for the treatment of prostate disease. The ESTs can be used in assays to
CC identify substances (e.g. triple helix forming oligonucleotides,
CC antisense oligonucleotides, ribozymes or antibodies) useful for treating
CC prostate disease, such as prostatitis and benign or malignant growth of
CC the prostate gland. This polynucleotide sequence represents the cDNA of a
CC urogenital sinus-derived expressed sequence tag of the invention.
XX
SQ Sequence 214 BP; 36 A; 68 C; 40 G; 70 T; 0 U; 0 Other;
Query Match 70.4%; Score 17.6; DB 3; Length 214;
Best Local Similarity 83.3%; Pred. No. 4.1e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AAAAAGCTGGAATCTCAGGCTGAGA 25
||||| ||||| ||||| ||||| |||||
Db 58 AATCACTGGAGGCTCAGGCTGAGA 35

RESULT 14
ADP28801/c
ID ADP28801 standard; DNA; 276 BP.
XX
XX AC ADP28801;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human secreted protein encoding sequence SEQ ID #799.
XX
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KW cancer; inflammatory; immune; ds; human secreted protein.
XX
XX Homo sapiens.
XX
XX WO2004035732-A2.
XX
XX 29-APR-2004.
XX
XX 28-AUG-2003; 2003WO-US026780.
XX
XX 29-AUG-2002; 2002US-0406576P.
XX 29-AUG-2002; 2002US-0406579P.
XX 29-AUG-2002; 2002US-0406585P.
XX 29-AUG-2002; 2002US-0406588P.
XX 29-AUG-2002; 2002US-0406609P.
XX 29-AUG-2002; 2002US-0406611P.
XX 29-AUG-2002; 2002US-0406612P.
XX 29-AUG-2002; 2002US-0406616P.
XX 29-AUG-2002; 2002US-0406640P.
XX 29-AUG-2002; 2002US-0406642P.
XX 29-AUG-2002; 2002US-0406646P.
XX 29-AUG-2002; 2002US-0406653P.
XX 29-AUG-2002; 2002US-0406655P.
XX 29-AUG-2002; 2002US-0406666P.
XX 17-SEP-2002; 2002US-0410944P.
XX 17-SEP-2002; 2002US-0410947P.
XX 17-SEP-2002; 2002US-0410948P.
XX 17-SEP-2002; 2002US-0410949P.
XX 17-SEP-2002; 2002US-0410953P.
XX 17-SEP-2002; 2002US-0410957P.
XX 17-SEP-2002; 2002US-0410958P.
XX 17-SEP-2002; 2002US-0410959P.
XX 17-SEP-2002; 2002US-0410960P.
XX 17-SEP-2002; 2002US-0410961P.
XX 17-SEP-2002; 2002US-0410962P.
XX 17-SEP-2002; 2002US-0411019P.
XX 17-SEP-2002; 2002US-0411022P.
XX 17-SEP-2002; 2002US-0411023P.
XX 17-SEP-2002; 2002US-0411024P.
XX 17-SEP-2002; 2002US-0411032P.

PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 03-JUN-2003; 2003US-0476609P.
PR 03-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RE, Huang MM, Kothakota S, Haishan L, Linnemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX
XX WPI; 2004-348438/32.
XX
XX New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
XX Claim 1; SEQ ID NO 799; 428pp; English.
XX
XX The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC anti-inflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein encoding sequence. The
CC present sequence is available on WIPOMEB and is not in the specification.
XX
XX Sequence 276 BP; 95 A; 57 C; 45 G; 79 T; 0 U; 0 Other;
SQ
Query Match 70.4%; Score 17.6; DB 12; Length 276;
Best Local Similarity 83.3%; Pred. No. 4.3e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Qy 1 AAAAACTGGAATCTCAGGCTGAG 24
|||||
Db 232 AAAAACTGGAATCTCAGGCTGTG 209
|||||
RESULT 15
ACH68960/c
ID ACH68960 standard; DNA; 522 BP.
XX
AC ACH68960;
XX
DT 29-JUL-2004 (first entry)
XX
DE Human genome derived single exon probe #2155.
XX
KW Human; probe; ss; gene expression; single exon probe; microarray;
KW alternative splicing event; genomic alteration.
XX
OS Homo sapiens.
XX
PN US2003194704-A1.
XX
PD 16-OCT-2003.
XX
XX 03-APR-2002; 2002US-00029386.
XX
XX 03-APR-2002; 2002US-00029386.
XX (PENN/) PENN S G.
XX (RANK/) RANK D R.
XX (HANZ/) HANZEL D K.
XX
XX Penn SG, Rank DR, Hanzel DK;
XX
XX WPI; 2004-119264/12.
XX
XX New human genome-derived single exon nucleic acid probes useful for human
PT gene expression analysis, for identifying or characterizing alternative
PT splicing events, for assessing genomic alterations or as tools for
PT surveying tissues.
XX
XX Claim 15; SEQ ID NO 2155; 80pp; English.
XX
XX The invention relates to a nucleic acid probe for measuring human gene
CC expression, comprising any of the 27,400 fully defined nucleotide
CC sequences in the specification, or their complements or fragments, and
CC encoding at least 8 amino acids of any of the 6888 amino acid sequences
CC fully defined in the specification. The probe is a single exon probe that
CC hybridizes under high stringency conditions to a nucleic acid molecule
CC expressed in human cells or tissues. Also included are a spatially-
CC addressable set of single exon nucleic acid probes for measuring human
CC gene expression (comprising a plurality of single exon nucleic acid
CC probes cited above, where each of the plurality of probes is separately
CC and addressably isolatable or amplifiable from the plurality), a single
CC exon microarray for measuring human gene expression, a method of
CC measuring human gene expression, a vector comprising the single exon
CC probe cited above, an ORF-encoded peptide comprising at least 8
CC contiguous amino acids of any of the above-mentioned amino acid
CC sequences (optionally with conservative amino acid substitutions), an
CC isolated antibody that binds specifically to a peptide cited above,
CC methods of selling and/or licensing single exon probes or microarrays to
CC a customer desiring to measure gene expression, a method of providing
CC human gene expression data by subscription, and a computer-readable
CC storage medium which contains a database having a plurality of records
CC (each record including data on the expression of a single exon probe
CC cited above. The probe, methods and apparatus are useful in gene
CC expression analysis. The probes may be used as tools for surveying
CC tissues to detect the presence of expressed messages that contain their
CC specific exon, or in constructing genome-derived single exon microarrays.
CC In addition, the probes are used in identifying and characterising
CC alternative splicing events, in detecting and characterising gross
CC alterations in the genomic locus that includes their exon, in assessing
CC smaller genomic alterations, in priming the synthesis of nucleic acids,
CC or in expressing the ORF-encoded peptide. The present sequence is a human
CC single exon probe of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20030194704

XX SQ Sequence 522 BP; 153 A; 126 C; 68 G; 175 T; 0 U; 0 Other;
Query Match 70.4%; Score 17.6; DB 12; Length 522;
Best Local Similarity 83.3%; Pred. NO. 4.8e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 AAAAAGCTGGAATCTCAGGCTGAG 24
Db 264 AAAAAGCTGGAATGTGAAGGCTGTG 241
RESULT 16
AAH09027
ID AAH09027 standard; cDNA; 561 BP.
XX AC AAH09027;
XX DT 26-JUN-2001 (first entry)
XX DE Human cDNA clone (3'-primer) SEQ ID NO:5862.
XX KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX OS Homo sapiens.
XX PN EP1074617-A2.
XX PD 07-FEB-2001.
XX PF 28-JUL-2000; 2000EP-00116126.
XX PR 29-JUL-1999; 99JP-00248036.
XX PR 27-AUG-1999; 99JP-00300253.
XX PR 11-JAN-2000; 2000JP-00118776.
XX PR 02-MAY-2000; 2000JP-00183767.
XX PR 09-JUN-2000; 2000JP-00241899.
XX (HELI-) HELIX RES INST.
XX Ota T, Isogai T, Nishikawa T, Hayaashi K, Saito K, Yamamoto J;
XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WPI; 2001-318749/34.
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.
XX
PS Claim 3; SEQ ID NO 5862; 2537pp + Sequence Listing; English.
XX
CC The present invention describes primer sets for synthesising 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesising polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the

CC present invention
XX SQ Sequence 561 BP; 163 A; 134 C; 121 G; 137 T; 0 U; 6 Other;
Query Match 70.4%; Score 17.6; DB 4; Length 561;
Best Local Similarity 83.3%; Pred. NO. 4.8e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 AAAAAGCTGGAATCTCAGGCTGAG 24
Db 247 AAAGCACTGGAATTACAGGCTGAG 270
RESULT 17
ACN39840
ID ACN39840 standard; cDNA; 1266 BP.
XX AC ACN39840;
XX DT 18-NOV-2004 (first entry)
XX DE Tumour-associated antigenic target (TAT) cDNA DNA326021, SEQ ID NO:4237.
XX KW Tumour-associated antigenic target; TAT; human; overexpression; cancer; tumour; diagnosis; cell proliferative disorder; breast cancer; colorectal cancer; lung cancer; ovarian cancer; liver cancer; central nervous system cancer; bladder cancer; pancreatic cancer; cervical cancer; melanoma; leukaemia; hybridisation probe; chromosome identification; chromosome mapping; gene mapping; gene therapy; cytostatic; gene; ss.
XX OS Homo sapiens.
XX PN WO2004030615-A2.
XX PD 15-APR-2004.
XX PF 29-SEP-2003; 2003WO-US028547.
XX PR 02-OCT-2002; 2002US-0414971P.
XX PA (GETH) GENENTECH INC.
XX PI Wu TD, Zhang Z, Zhou Y;
XX WPI; 2004-347921/32.
XX DR P-PSDB; ABM81641.
XX PT New tumor-associated antigenic target polypeptides and nucleic acids, useful in preparing a medicament for treating or detecting a proliferative disorder, e.g. breast, lung, colorectal, ovarian or prostate cancer or tumor.
XX
PS Claim 1; SEQ ID NO 4237; 7273pp; English.
XX
CC The invention relates to human tumour-associated antigenic target (TAT) polypeptides, and their related nucleic acids. The TAT polypeptides are overexpressed in cancer tissues compared to normal tissues, and may thus serve as effective targets for the diagnosis and treatment of cancer in mammals. The invention also relates to nucleic acid and polypeptide sequences at least 80% identical to the TAT nucleic acids and polypeptides; expression vectors and host cells comprising a TAT nucleic acid; an antibody specific for a TAT polypeptide; a peptide or organic molecule which binds to a TAT polypeptide; fusion proteins comprising a TAT polypeptide; and methods and compositions for the treatment or diagnosis of cancer in mammals. TAT polypeptides, nucleic acids, antibodies, antagonists, binding molecules and compositions are useful for diagnosing or treating a cell proliferative disorder associated with increased TAT expression, particularly cancers such as breast cancer, colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder cancer, pancreatic cancer, cervical cancer, cancers of the central nervous system, melanoma and leukaemia. TAT nucleic acids may further be used as hybridisation probes, in chromosome and gene mapping, in

CC chromosome identification and in gene therapy. The present sequence
CC represents a TAT nucleic acid of the invention

XX
SQ Sequence 1266 BP; 432 A; 210 C; 302 G; 322 T; 0 U; 0 Other;

Query Match	Score 17.6;	DB 13;	Length 1266;
Best Local Similarity	83.3%	Pred. No. 5.5e+02;	
Matches 20;	Conservative	0;	Mismatches 4;
			Indels 0;
			Gaps 0;

Query Match 70.4%; Score 17.6; DB 13; Length 1266;

Best Local Similarity 83.3%; Pred. No. 5.5e+02;
 Matches 20: Conservative 0: Mismatches 4: Indels 0: Gaps 0:

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

QY
1 AAAA AACTGGAATCTCAGGCTGAG 24

D_b 883 AAAAAGCTGAAATGTAAGGCTGTG 906

RESULT 18

AAH13739/C

ID AAH13739 standard; cDNA; 1688 BP.

XX

AC AAH13739;

XX

DT 26-JUN-2001 (first entry)

100-10641
 DEC 19 1964
 U.S. DEPT. OF COMMERCE
 NATIONAL BUREAU OF STANDARDS-10641

DE
yy
Human cDNA sequence SEQ ID NO:10641.

Human: primer

**XX
XX
MY
TRANSPORT**

Homo sapiens.

XX
33
1000-2430 000000

PN EP1074617-A2.

XX

PD 07-FEB-2001.

XX

PF 28-JUL-2000; 2000EP-001



PR 29-JUL-1999; 99JP-00248036.

PR 27-AUG-1999; 99JP-

PR	11-JAN-2000; 2000JP-00118776.	
PR	PR	02-MAY-2000; 2000JP-00183767.
PR	PR	09-JUN-2000; 2000JP-00241899.
XX	XX	
XX		(HELI-) HELIX RES INST.
XX	PA	
XX	PI	Ota T, Isoqai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI	PI	Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX	XX	
XX	DR	WPI; 2001-318749/34.
XX	XX	
PT	PT	Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.
PT	PT	

PS Claim 8: SEQ ID NO 10641: 2537bp + Sequence Listing: English.

The present invention describes primer sets for synthesising 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dr primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesising polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAH92440 to AAH95893

CC used as hybridisation probes, in chromosome and gene mapping, in
CC chromosome identification and in gene therapy. The present sequence
CC represents a TAT nucleic acid of the invention

XX
SQ Sequence 2177 BP; 703 A; 363 C; 442 G; 669 T; 0 U; 0 Other;
Query Match 70.4%; Score 17.6; DB 13; Length 2177;
Best Local Similarity 83.3%; Pred. No. 6e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAACTGGAATCTCAGGCTGAG 24
|||||
Db 750 AAAAACTGGAATGTAAGGCTGTG 773

RESULT 20
ADRO7065/c
ID ADRO7065 standard; cDNA; 3081 BP.
XX AC ADR07065;
XX
XX 04-NOV-2004 (first entry)
DE Full length human cDNA useful for treating neurological disease Seq 571.
XX
XX Gene; ss; human; oligo-capping method; diagnostic marker; gene therapy;
KW osteoporosis; neurological disease; Alzheimer's disease;
KW Parkinson's disease; dementia; short memory; cancer;
KW sense or motor function; emotional reaction; fear response; panic;
KW osteopathic; neuroprotective; nootropic; antiparkinsonian; cytostatic;
KW tranquiliser.
XX
XX Homo sapiens.
XX
XX EP1447413-A2.
XX 18-AUG-2004.
XX
XX 12-FEB-2004; 2004EP-00003145.
XX
XX 14-FEB-2003; 2003JP-00102207.
PR 09-MAY-2003; 2003JP-00131452.
XX
XX (REAS-) RES ASSOC BIOTECHNOLOGY.
PA
XX
XX Isogai T, Yamamoto J, Nishikawa T, Isono Y, Sugiyama T, Otsuki T;
PI Wakamatsu A, Ishii S, Nagai K, Irie R;
XX
XX WPI; 2004-583265/57.
DR
DR P-PSDB; ADRO9021.
XX
XX New 1995 cDNA, useful for treating osteoporosis, neurological diseases,
PT Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.
PT
PS Claim 1; SEQ ID NO 571; 2686pp; English.
XX
XX This invention relates to novel, isolated full length human cDNA
CC molecules and the encoded proteins thereof. Specifically, it refers to
CC cDNA clones obtained by an oligo-capping method, where none of these
CC clones are identical to any known human mRNAs. The present invention
CC describes an immunoassay to identify agonists and antagonists, as well as
CC antibodies, antisense molecules and siRNAs that can all be used to bind
CC to and modulate expression of the cDNA molecules. As such, these
CC molecules are useful for diagnostic markers or therapeutic targets for
CC the various diseases or morbid states. In particular, they are useful in
CC gene therapy for treating osteoporosis, neurological disease, Alzheimer's
CC disease, Parkinson's disease, dementia, short memory and various cancers,
CC as well as for maintaining equilibrium of sense or motor function, and
CC for treating emotional reaction, fear response and panic. Accordingly,
CC they exhibit osteopathic, neuroprotective, nootropic, antiparkinsonian,
CC cytotstatic and tranquiliser activities. This polynucleotide is a full
CC length human cDNA sequence of the invention. NOTE: This sequence is not
CC given in the sequence listing of the specification but can be obtained on

CC CD-ROM from the European Patent Office, Vienna Sub-office.
XX
SQ Sequence 3081 BP; 861 A; 613 C; 682 G; 925 T; 0 U; 0 Other;
Query Match 70.4%; Score 17.6; DB 13; Length 3081;
Best Local Similarity 83.3%; Pred. No. 6.4e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AAAAACTGGAATCTCAGGCTGAGA 25
|||||
Db 401 AAAAACTGGAATCAGGCTGACA 378

RESULT 21
ABX63354/c
ID ABX63354 standard; cDNA; 5271 BP.
XX AC ABX63354;
XX
XX 25-FEB-2003 (first entry)
DE Human cDNA #354 differentially expressed in activated vascular tissue.
XX
XX Human; gene; ss; vascular tissue; cytostatic; atherosclerosis; cardiant;
KW hypotensive; antidiabetic; gynaecological; vasotropic; cerebroprotective;
KW gene therapy; vascular disease; cancer; coronary; artery disease;
KW hypertension; diabetes; pre-eclampsia; restenosis;
KW ischaemia-reperfusion injury; stroke.
XX
XX Homo sapiens.
XX
XX US2002137081-A1.
XX
XX 26-SEP-2002.
XX
XX 08-JAN-2002; 2002US-00044090.
XX
XX 28-JUL-2000; 2000US-0222469P.
PR 08-JAN-2001; 2001US-0260483P.
XX
XX (BAND/) BANDMAN O.
XX Bandman O;
XX
XX WPI; 2003-110597/10.
DR
XX
XX Combination for diagnosing, staging, treating, or monitoring the
PT progression of treatment of a vascular disease, e.g. atherosclerosis,
PT comprises several cDNAs that are differentially expressed in activated
PT vascular tissue.
XX
XX Claim 1; Page; 18pp; English.
PS
XX This invention relates to a combination comprising several cDNAs that are
CC differentially expressed in activated vascular tissue. The invention also
CC discloses a high throughput method for detecting differentially expressed
CC cDNAs in a sample. The cDNAs of the invention may have
CC antiarteriosclerotic; cytostatic; cardiant; hypotensive; antidiabetic;
CC gynaecological; vasotropic and cerebroprotective activities and may be
CC used in gene therapy. The cDNAs of the invention may be used in a high-
CC throughput methods for detecting differential expression of one or more
CC cDNAs in a sample, or screening several molecules or compounds to
CC identify a molecule or compound that specifically binds a cDNA of the
CC invention. A protein encoded by the cDNA may be used to screen several
CC molecules or compounds to identify a ligand that specifically binds to
CC the protein, or to produce or purify an antibody to the protein that can
CC be used to detect a protein in a sample or purify a natural or
CC recombinant protein from a sample. The nucleotides may be useful for
CC diagnosing, staging, treating, or monitoring the progression of treatment
CC of a vascular disease, e.g. atherosclerosis, cancer, coronary artery
CC disease, hypertension, diabetes, pre-eclampsia, ischaemia-reperfusion
CC injury, restenosis, or stroke. The cDNAs can also be used for large-scale
CC genetic or gene expression analysis of several new nucleic acid

CC molecules. Antibodies to the proteins encoded by the cDNAs are useful for
CC diagnosing pre-pathologic disorders, and chronic or acute diseases
CC associated with abnormalities in the expression, amount or distribution
CC of the protein. The present sequence represents a cDNA of the invention
CC that is differentially expressed in activated vascular tissue. Note: The
CC sequence data for this patent did not form part of the specification, but
CC was obtained in electronic format directly from USPTO at
CC <http://seqdata.uspto.gov/sequence.html?DocID=20020137081>
XX
SQ Sequence 5271 BP; 1374 A; 1122 C; 1141 G; 1294 T; 0 U; 340 Other;

Query Match 70.4%; Score 17.6; DB 8; Length 5271;
Best Local Similarity 83.3%; Pred. No. 6.9e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAACTGGATCTCAGGCTGAG 24
|||||
Db 3345 AAAAATACTGGATTACAGGCTGAG 3322

RESULT 22
ADJ56534/c
ID ADJ56534 standard; cDNA; 5271 BP.

AC ADJ56534;

XX 06-MAY-2004 (first entry)

DT Rat cDNA differentially expressed in MYCN activated cells SeqID 340.

DE rat; differential expression; transactivator; proto-oncogene;
KW neuroblastoma; small cell lung cancer; cytostatic; gene therapy; ss;
KW MYCN activated cell.

XX Rattus norvegicus.

OS US2003119009-A1.

PN 26-JUN-2003.

XX 25-FEB-2002; 2002US-00084817.

XX 23-FEB-2001; 2001US-0270784P.

XX (STUA/) STUART S G.

PA (NUCH/) NUCHTERN J G.

PA (PLON/) PLON S E.

PA (SHOH/) SHOHET J M.

XX Stuart SG, Nuchtern JG, Plon SE, Shohet JM;

PI WPI; 2003-635698/60.

XX New genes regulated by MYCN activation, useful in gene therapy.

PT Particularly for treating a subject with e.g. neuroblastoma or other
PT cancers, or for diagnosing, staging or monitoring the treatment of the
PT cancer.

XX Claim 1; SEQ ID NO 340; 27pp; English.

XX This invention relates to novel isolated cDNAs that are differentially
CC expressed in MYCN activated cells. Specifically, it refers to
CC polynucleotide sequences that exhibit differential expression patterns in
CC cells activated by the transactivator MYCN, where MYCN is a proto-
CC oncogene that is amplified in neuroblastoma cells and is common in small
CC cell lung cancers. The present invention describes these cDNA molecules
CC as useful for in hybridisation assays to detect expression of nucleic
CC acids (or complementary nucleic acids) in a present in a given sample, as
CC well as for screening assays by identifying molecules or compounds that
CC specifically bind the cDNA as a ligand and modulate function or activity.
CC Accordingly, these compositions exhibit cytostatic activity and can also
CC be used for gene therapy purposes. This polynucleotide sequence is a cDNA
CC that is differentially expressed in MYCN activated cells, given in an

CC exemplification of the invention. NOTE: This sequence does not appear in
CC the printed specification but has been obtained in electronic format from
CC the US Patent Office at
CC <ftp://seqdata.uspto.gov/sequence.html?DocID=20030119009>.

XX Sequence 5271 BP; 1374 A; 1122 C; 1141 G; 1294 T; 0 U; 340 Other;

Query Match 70.4%; Score 17.6; DB 10; Length 5271;
Best Local Similarity 83.3%; Pred. No. 6.9e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAACTGGATCTCAGGCTGAG 24
|||||
Db 3345 AAAAATACTGGATTACAGGCTGAG 3322

RESULT 23
ABS76688
ID ABS76688 standard; DNA; 14815 BP.

XX ABS76688;

XX 12-DEC-2002 (first entry)

DE PRRS virus DNA sequence.

XX Porcine reproductive and respiratory syndrome; PRRS; virus;

XX open reading frame; ORF; ORF1; ORF2; ORF3; ORF4; ORF5; ORF6; ORF7;

XX PRRS infection; virucide; vaccine; pig; ds.

XX Porcine reproductive and respiratory syndrome virus.

XX WO200272802-A2.

XX 19-SEP-2002.

XX 07-MAR-2002; 2002WO-EP002486.

XX 09-MAR-2001; 2001US-0274603P.

XX (BOEH) BOEHRINGER INGELHEIM VETMEDICA GMBH.

XX Elbers K, Peschke S, Schuetz B;

XX WPI; 2002-723343/78.

XX New attenuated European porcine reproductive and respiratory syndrome
PT (PRRS) virus, useful for manufacturing a vaccine against PRRS, i.e. for
PT the prophylaxis and treatment of PRRS infection.

XX Claim 6; Page 45-50; 75pp; English.

XX The present invention relates to a new attenuated European porcine
CC reproductive and respiratory syndrome (PRRS) virus encoded by a nucleic
CC acid comprising open reading frame (ORF)1, ORF2, ORF3, ORF4, ORF5, ORF6
CC and ORF7. The attenuated European PRRS virus is useful in the manufacture
CC of a vaccine against PRRS, i.e. for the prophylaxis and treatment of PRRS
CC infection. The present nucleic acid sequence represents the PRRS virus
CC sequence of the invention

XX Sequence 14815 BP; 3184 A; 4082 C; 3717 G; 3830 T; 0 U; 2 Other;

Query Match 70.4%; Score 17.6; DB 6; Length 14815;
Best Local Similarity 83.3%; Pred. No. 8.2e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AAAAACTGGATCTCAGGCTGAGA 25
|||||

Db 9965 AAAAATCTGAATTAAGGCTAGGA 9988

RESULT 24
AD213631_4/c

```
Continuation (5 of 5) of ADZ13631 from base 40001 (Human cancer-associated genomic DNA
WP Sequence split into 5 fragments LOCUS ADZ13631 Accession Adz13631
WP Fragment Name Begin End
WP ADZ13631_0 1 110000
WP ADZ13631_1 100001 210000
WP ADZ13631_2 200001 310000
WP ADZ13631_3 300001 410000
WP ADZ13631_4 400001 420555

Query Match 70.4%; Score 17.6; DB 14; Length 20555;
Best Local Similarity 83.3%; Pred. No. 8.6e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AAAAACTGGAATCTCAGGCTGAGA 25
||||| ||| ||||| |||
Db 18583 AAAAACTGAAAAATCAGGCTTAGA 18560

RESULT 25
ADZ13620_4/c
Continuation (5 of 5) of ADZ13620 from base 40001 (Human cancer-associated genomic DNA
WP Sequence split into 5 fragments LOCUS ADZ13620 Accession Adz13620
WP Fragment Name Begin End
WP ADZ13620_0 1 110000
WP ADZ13620_1 100001 210000
WP ADZ13620_2 200001 310000
WP ADZ13620_3 300001 410000
WP ADZ13620_4 400001 420555

Query Match 70.4%; Score 17.6; DB 14; Length 20555;
Best Local Similarity 83.3%; Pred. No. 8.6e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AAAAACTGGAATCTCAGGCTGAGA 25
||||| ||| ||||| |||
Db 18583 AAAAACTGAAAAATCAGGCTTAGA 18560

RESULT 26
AAD43379/c
ID AAD43379 standard; DNA; 24526 BP.
XX
AC AAD43379;
XX
DT 13-DEC-2002 (first entry)
XX
DE Human transporter gene.
XX
KW Human; transporter protein; transgenic; pharmacogenomic analysis;
KW gene therapy; chromosome 1; gene; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT variation replace(48,G)
FT /tag= a
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT variation replace(132,A)
FT /tag= b
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT variation replace(724,C)
FT /tag= c
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT variation replace(1558,G)
FT /tag= d
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT variation replace(1577,G)
FT /tag= e
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT CDS 1997..22526
FT /tag= f
FT /product= "Human transporter protein"
FT exon 1997..2121
```

```
FT intron /tag= g
FT 2122..4732
FT /tag= h
FT variation replace(2487,A)
FT /tag= i
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT variation replace(2634,C)
FT /tag= j
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT variation replace(4352,G)
FT /tag= k
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT exon 4733..4872
FT /tag= l
FT intron 4873..5004
FT /tag= m
FT exon 5005..5115
FT /tag= n
FT intron 5116..5781
FT /tag= o
FT variation replace(5157,C)
FT /tag= p
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT variation replace(5658,T)
FT /tag= q
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT exon 5782..5957
FT /tag= r
FT variation replace(5945,C)
FT /tag= s
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT intron 5958..7770
FT /tag= t
FT variation replace(6281,C)
FT /tag= u
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT variation replace(6452,C)
FT /tag= v
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT variation replace(6610,T)
FT /tag= w
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT variation replace(7247,C)
FT /tag= x
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT variation replace(7360,G)
FT /tag= y
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT variation replace(7644,T)
FT /tag= z
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT exon 7771..7935
FT /tag= aa
FT intron 7936..8470
FT /tag= ab
FT variation replace(8127,C)
FT /tag= ac
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT variation replace(8317,A)
FT /tag= ad
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT exon 8471..8623
FT /tag= ae
FT intron 8624..8917
FT /tag= af
FT exon 8918..9000
FT /tag= ag
FT intron 9001..9777
FT /tag= ah
FT variation replace(9079,A)
FT /tag= ai
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT variation replace(9537,G)
```



```
FT      /*tag= j  
FT      /standard_name= "Single nucleotide polymorphism (SNP) "  
FT      replace(4352,G)  
FT      /*tag= k  
FT      /standard_name= "Single nucleotide polymorphism (SNP) "  
FT      4733. .4872  
FT      /*tag= l  
FT      4873. .5004  
FT      /*tag= m  
FT      5005. .5115  
FT      /*tag= n  
FT      5116. .5781  
FT      /*tag= o  
FT      replace(5157,C)  
FT      /*tag= p  
FT      /standard_name= "Single nucleotide polymorphism (SNP) "  
FT      replace(5658,T)  
FT      /*tag= q  
FT      /standard_name= "Single nucleotide polymorphism (SNP) "  
FT      5782. .5957  
FT      /*tag= r  
FT      replace(5945,C)  
FT      /*tag= s  
FT      /standard_name= "Single nucleotide polymorphism (SNP) "  
FT      5958. .7770  
FT      /*tag= t  
FT      replace(6281,C)  
FT      /*tag= u  
FT      /standard_name= "Single nucleotide polymorphism (SNP) "  
FT      replace(6452,C)  
FT      /*tag= v  
FT      /standard_name= "Single nucleotide polymorphism (SNP) "  
FT      replace(6610,T)  
FT      /*tag= w  
FT      /standard_name= "Single nucleotide polymorphism (SNP) "  
FT      replace(7247,C)  
FT      /*tag= x  
FT      /standard_name= "Single nucleotide polymorphism (SNP) "  
FT      replace(7360,G)  
FT      /*tag= y  
FT      /standard_name= "Single nucleotide polymorphism (SNP) "  
FT      replace(7644,T)  
FT      /*tag= z  
FT      /standard_name= "Single nucleotide polymorphism (SNP) "  
FT      7771. .7935  
FT      /*tag= aa  
FT      7936. .8470  
FT      /*tag= ab  
FT      replace(8127,C)  
FT      /*tag= ac  
FT      /standard_name= "Single nucleotide polymorphism (SNP) "  
FT      replace(8317,A)  
FT      /*tag= ad  
FT      /standard_name= "Single nucleotide polymorphism (SNP) "  
FT      8471. .8623  
FT      /*tag= ae  
FT      8624. .8917  
FT      /*tag= af  
FT      8918. .9000  
FT      /*tag= ag  
FT      9001. .9777  
FT      /*tag= ah  
FT      replace(9079,A)  
FT      /*tag= ai  
FT      /standard_name= "Single nucleotide polymorphism (SNP) "  
FT      replace(9537,G)  
FT      /*tag= aj  
FT      /standard_name= "Single nucleotide polymorphism (SNP) "  
FT      9778. .9925  
FT      /*tag= ak  
FT      9926. .10221  
FT      /*tag= al  
FT      10222. .10335
```

```
FT      intron  
FT      10336. .10539  
FT      /*tag= an  
FT      10540. .10617  
FT      /*tag= ao  
FT      10618. .11197  
FT      /*tag= ap  
FT      11198. .11293  
FT      /*tag= aq  
FT      11294. .13338  
FT      /*tag= ar  
FT      replace(12302,G)  
FT      /*tag= as  
FT      /standard_name= "Single nucleotide polymorphism (SNP) "  
FT      replace(12354,C)  
FT      /*tag= at  
FT      /standard_name= "Single nucleotide polymorphism (SNP) "  
FT      replace(12487,C)  
FT      /*tag= au  
FT      /standard_name= "Single nucleotide polymorphism (SNP) "  
FT      replace(13197. .13199,CA)  
FT      /*tag= av  
FT      /standard_name= "Single nucleotide polymorphism (SNP) "  
FT      replace(13257,G)  
FT      /*tag= aw  
FT      /standard_name= "Single nucleotide polymorphism (SNP) "  
FT      13339. .13445  
FT      /*tag= ax  
FT      13446. .14214  
FT      /*tag= ay  
FT      14215. .14284  
FT      /*tag= az  
FT      14285. .14400  
FT      /*tag= ba  
FT      14401. .14493  
FT      /*tag= bb  
FT      14494. .15980  
FT      /*tag= bc  
FT      replace(14541,A)  
FT      /*tag= bd  
FT      /standard_name= "Single nucleotide polymorphism (SNP) "  
FT      replace(14545,G)  
FT      /*tag= be  
FT      /standard_name= "Single nucleotide polymorphism (SNP) "  
FT      replace(15041,A)  
FT      /*tag= bf  
FT      /standard_name= "Single nucleotide polymorphism (SNP) "  
FT      replace(15053,C)  
FT      /*tag= bg  
FT      /standard_name= "Single nucleotide polymorphism (SNP) "  
FT      replace(15065,G)  
FT      /*tag= bh  
FT      /standard_name= "Single nucleotide polymorphism (SNP) "  
FT      replace(15108,C)  
FT      /*tag= bi  
FT      /standard_name= "Single nucleotide polymorphism (SNP) "  
FT      15981. .16262  
FT      /*tag= bj  
FT      16263. .17597  
FT      /*tag= bk  
FT      replace(16273. .16275,CG)  
FT      /*tag= bl  
FT      /standard_name= "Single nucleotide polymorphism (SNP) "  
FT      replace(17424,T)  
FT      /*tag= bm  
FT      /standard_name= "Single nucleotide polymorphism (SNP) "  
FT      17598. .17652  
FT      /*tag= bn  
FT      replace(17627,A)  
FT      /*tag= bo  
FT      /standard_name= "Single nucleotide polymorphism (SNP) "  
FT      17653. .18842  
FT      /*tag= bp
```



```
FT variation replace(18427,C)
FT /tag= bq
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(18813,G)
FT /tag= br
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT 18843..18988
FT /tag= bs
FT 18989..20477
FT /tag= bt
FT replace(19035,C)
FT /tag= bu
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(19182,C)
FT
Query Match 70.4%; Score 17.6; DB 13; Length 24526;
Best Local Similarity 83.3%; Pred. No. 8.9e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 2 AAAAACTGGAATCTCAGGCTGAGA 25
Db 3296 AAGAACTGGAAACAGAGGCTGAGA 3273
RESULT 28
AB224431/C
ID AB224431 standard; cDNA; 26320 BP.
XX AC
XX AB224431;
XX
DT 18-MAR-2003 (first entry)
DE Human constitutively active receptor (hCAR) gene.
XX
KW Constitutively active reactor; human; hCAR; receptor;
KW G-protein coupled receptor; cytotostatic; neuroleptic; antiparkinsonian;
KW single nucleotide polymorphism; SNP; chromosome 4; gene therapy; gene;
KW db.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT variation replace(469,C)
FT /tag= h
FT /standard_name= "Single nucleotide polymorphism"
FT /note= "SNP CV1221921, position 8317095 on chromosome 4"
FT replace(469,C)
FT /tag= g
FT /standard_name= "Single nucleotide polymorphism"
FT /note= "SNP CV1221921, position 8317095 on chromosome 4"
FT replace(568,A)
FT /tag= i
FT /standard_name= "Single nucleotide polymorphism"
FT /note= "SNP CV1221920, position 8317194 on chromosome 4"
FT replace(865,C)
FT /tag= j
FT /standard_name= "Single nucleotide polymorphism"
FT /note= "SNP CV1221919, position 8317491 on chromosome 4"
FT replace(1046,T)
FT /tag= k
FT /standard_name= "Single nucleotide polymorphism"
FT /note= "SNP CV1221918, position 8317672 on chromosome 4"
FT replace(1637,T)
FT /tag= l
FT /standard_name= "Single nucleotide polymorphism"
FT /note= "SNP CV1221917, position 8318263 on chromosome 4"
FT replace(2349,G)
FT /tag= m
FT /standard_name= "Single nucleotide polymorphism"
FT /note= "SNP CV1221916, position 8319575 on chromosome 4"
FT replace(3335,C)
FT /tag= n
FT /standard_name= "Single nucleotide polymorphism"
```

```
FT variation
FT /tag= o
FT /standard_name= "Single nucleotide polymorphism"
FT /note= "SNP CV7662683, position 8321056 on chromosome 4"
FT replace(6658,T)
FT /tag= p
FT /standard_name= "Single nucleotide polymorphism"
FT /note= "SNP CV1221914, position 8323284 on chromosome 4"
FT replace(8080,C)
FT /tag= q
FT /standard_name= "Single nucleotide polymorphism"
FT /note= "SNP CV1221913, position 8324706 on chromosome 4"
FT replace(8086,C)
FT /tag= r
FT /standard_name= "Single nucleotide polymorphism"
FT /note= "SNP CV1221912, position 8324712 on chromosome 4"
FT replace(8090,C)
FT /tag= s
FT /standard_name= "Single nucleotide polymorphism"
FT /note= "SNP CV1221911, position 8324716 on chromosome 4"
FT replace(8627,G)
FT /tag= t
FT /standard_name= "Single nucleotide polymorphism"
FT /note= "SNP CV1221910, position 8325253 on chromosome 4"
FT replace(9804,G)
FT /tag= u
FT /standard_name= "Single nucleotide polymorphism"
FT /note= "SNP CV1221909, position 8326430 on chromosome 4"
FT 10001..16380
FT /tag= a
FT /product= "Human constitutively active receptor"
FT /note= "contains introns"
FT 10001..10668
FT /tag= b
FT 10669..11548
FT /tag= c
FT 11549..11662
FT /tag= d
FT 11663..16073
FT /tag= e
FT replace(13860,T)
FT /tag= v
FT /standard_name= "Single nucleotide polymorphism"
FT /note= "SNP CV1221908, position 8330486 on chromosome 4"
FT replace(13889,G)
FT /tag= w
FT /standard_name= "Single nucleotide polymorphism"
FT /note= "SNP CV1221907, position 8330515 on chromosome 4"
FT replace(14646,G)
FT /tag= x
FT /standard_name= "Single nucleotide polymorphism"
FT /note= "SNP CV1221906, position 8331272 on chromosome 4"
FT replace(14758,G)
FT /tag= y
FT /standard_name= "Single nucleotide polymorphism"
FT /note= "SNP CV1221905, position 8331384 on chromosome 4"
FT replace(15134,T)
FT /tag= z
FT /standard_name= "Single nucleotide polymorphism"
FT /note= "SNP CV1221904, position 8331760 on chromosome 4"
FT replace(15208,-)
FT /tag= aa
FT /standard_name= "Single nucleotide polymorphism"
FT /note= "SNP CV7664553, position 8331834 on chromosome 4"
FT replace(15253,G)
FT /tag= ab
FT /standard_name= "Single nucleotide polymorphism"
FT /note= "SNP CV1221902, position 8331879 on chromosome 4"
FT 16074..16383
FT /tag= f
FT replace(16279,A)
FT /tag= ac
```

FT /standard_name= "Single nucleotide polymorphism"
FT /note= "SNP CV1221901, position 8332905 on chromosome 4"
FT replace(16448,T)
FT /*tag= ad
FT
FT /standard_name= "Single nucleotide polymorphism"
FT /note= "SNP CV1221900, position 8333074 on chromosome 4"
FT replace(16499,G)
FT /*tag= ae
FT /standard_name= "Single nucleotide polymorphism"
FT /note= "SNP CV1221899, position 8333125 on chromosome 4"
FT replace(16534,T)
FT /*tag= af
FT /standard_name= "Single nucleotide polymorphism"
FT /note= "SNP CV1221898, position 8333160 on chromosome 4"
FT replace(19520,G)
FT /*tag= ag
FT /standard_name= "Single nucleotide polymorphism"
FT /note= "SNP CV1221897, position 8336146 on chromosome 4"
FT replace(19647,C)
FT /*tag= ah
FT /standard_name= "Single nucleotide polymorphism"
FT /note= "SNP CV1221896, position 8336273 on chromosome 4"
FT replace(19999,T)
FT /*tag= ai
FT /standard_name= "Single nucleotide polymorphism"
FT /note= "SNP CV1221895, position 8336625 on chromosome 4"
FT replace(20807,T)
FT /*tag= ak
FT /standard_name= "Single nucleotide polymorphism"
FT /note= "SNP CV1221893, position 8337433 on chromosome 4"
FT replace(23179,A)
FT /*tag= al
FT /standard_name= "Single nucleotide polymorphism"
FT /note= "SNP CV1221892, position 8339805 on chromosome 4"
FT replace(23224,T)
FT /*tag= am
FT /standard_name= "Single nucleotide polymorphism"
FT /note= "SNP CV1221891, position 8339850 on chromosome 4"
FT replace(23251,A)
FT /*tag= an
FT /standard_name= "Single nucleotide polymorphism"
FT /note= "SNP CV1221890, position 8339877 on chromosome 4"
FT replace(23666,-)
FT /*tag= ao
FT /standard_name= "Single nucleotide polymorphism"
FT /note= "SNP CV7664594, position 8340292 on chromosome 4"
FT replace(23671,G)
FT /*tag= ap
FT /standard_name= "Single nucleotide polymorphism"
FT /note= "SNP CV7664595, position 8340297 on chromosome 4"
FT replace(23770,A)
FT /*tag= aq
FT /standard_name= "Single nucleotide polymorphism"
FT /note= "SNP CV1221887, position 8340396 on chromosome 4"
FT replace(23795,-)
FT /*tag= ar
FT /standard_name= "Single nucleotide polymorphism"
FT /note= "SNP CV7664734, position 8340421 on chromosome 4"
FT replace(24037,A)
FT /*tag= as
FT /standard_name= "Single nucleotide polymorphism"
FT /note= "SNP CV1221885, position 8340663 on chromosome 4"
FT replace(24431,C)
FT /*tag= at
FT /standard_name= "Single nucleotide polymorphism"
FT /note= "SNP CV1221884, position 8341057 on chromosome 4"
FT replace(24556,C)
FT /*tag= au
FT /standard_name= "Single nucleotide polymorphism"
FT /note= "SNP CV1221883, position 8341182 on chromosome 4"
FT replace(29155,A)
FT /*tag= aj
FT /standard_name= "Single nucleotide polymorphism"

FT /note= "SNP CV1221894, position 8336781 on chromosome 4"
XX WO2002101005-A2.
XX
XX PD 19-DEC-2002.
XX
XX PF 30-MAY-2002; 2002WO-US016834.

Query Match 70.4%; Score 17.6; DB 10; Length 26320;
Best Local Similarity 83.3%; Pred. No. 9e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 AAAAAGCTGGAAATCTCAGGCTGAG 24
||||| ||||| ||||| ||||| ||
Db 4864 AAAACCTGGAAATGCCAGGCTAAG 4841

RESULT 29
ADZ59528/C
ID ADZ59528 standard; DNA; 37442 BP.
XX
XX AC ADZ59528;
XX
XX DT 30-JUN-2005 (first entry)
XX
DE Secondary hyperparathyroidism detection polymorphic gene IL12RB1, SEQ 22.
XX
XX KW secondary hyperparathyroidism; endocrine-gen.; antithyroid;
KW renal failure; nephrotropic; SNP detection;
KW single nucleotide polymorphism; SNP; gene; ds; IL12RB1.
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
FH variation replace(27442,A)
FT /*tag= a
FT /standard_name= "Single nucleotide polymorphism"
XX
XX JP2005102601-A.
XX
XX PD 21-APR-2005.
XX
XX PF 30-SEP-2003; 2003JP-00341015.
XX
XX PR 30-SEP-2003; 2003JP-00341015.
XX
XX PA (HYUB-) HYUBITTO GENOMICS KK.
XX (JIKE-) UNIV JIKEI.
XX
XX DR WPI; 2005-358641/37.
XX
XX PT Testing secondary hyperparathyroidism in chronic renal failure patient,
PT involves detecting variation in gene chosen from CACNA1C, CALCL, CH13L1,
PT EGF, FGF1, GFRA1, GPR56 and GPRK6.
XX
XX PS Claim 4; SEQ ID NO 22; 138pp; Japanese.
XX
XX CC The invention relates to a novel method for testing secondary
CC hyperparathyroidism in a chronic renal failure patient. The method
CC involves detecting a variation in a gene chosen from CACNA1C, CALCL,
CC CH13L1, EGF, FGF1, GFRA1, GPR56, GPRK6, IL10RA, IL12RB1, KCNJ14,
CC KCNQ1, ORCT14, PDGFRA, SCYB14, SLC12A1, SLC2A3, TGFBR3, TMEM1, CALCR,
CC IL17R, OSTF1, FGF6, HGF, MET, TGFBI and VEGF, or detecting the base in a
CC polymorphism region existing in the vicinity of any one of the genes. The
CC invention further comprises a reagent or kit for testing secondary
CC hyperparathyroidism in a chronic renal failure patient. This
CC polynucleotide sequence represents the polymorphism containing human
CC IL12RB1 gene of the invention.
XX
XX SQ Sequence 37442 BP; 9374 A; 9810 C; 9241 G; 9017 T; 0 U; 0 Other;

Query Match 70.4%; Score 17.6; DB 14; Length 37442;
Best Local Similarity 83.3%; Pred. No. 9.5e+02;

Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 AAAAACTGGAATCTCAGGCTGAG 24
|||||
Db 18487 AAAAATCTGGAATTACAGGCTGAG 18464

RESULT 30

ADZ59517/C
ID ADZ59517 standard; DNA; 37442 BP.

XX ADZ59517;

XX 30-JUN-2005 (first entry)

XX Secondary hyperparathyroidism detection polymorphic gene IL12RB1, SEQ 11.

XX secondary hyperparathyroidism; endocrine-gen.; antithyroid;

KW renal failure; nephrotropic; SNP detection;

KW single nucleotide polymorphism; SNP; gene; ds; IL12RB1.

XX Homo sapiens.

XX Key Location/Qualifiers

FT variation replace(10001,G)

FT /*tag= a

FT /standard_name= "Single nucleotide polymorphism"

XX JP2005102601-A.

XX 21-APR-2005.

XX 30-SEP-2003; 2003JP-00341015.

XX 30-SEP-2003; 2003JP-00341015.

XX (HYUB-) HYUBITTO GENOMICS KK.

XX (JIK2-) UNIV JIKEI.

XX WPI; 2005-358641/37.

XX Testing secondary hyperparathyroidism in chronic renal failure patient,
PT involves detecting variation in gene chosen from CACNA1C, CALCL, CHI3L1,
PT EGF, FGF1, GFRA1, GPR56 and GPRK6.

XX Claim 4; SEQ ID NO 11; 138pp; Japanese.

XX The invention relates to a novel method for testing secondary
CC hyperparathyroidism in a chronic renal failure patient. The method
CC involves detecting a variation in a gene chosen from CACNA1C, CALCL,
CC CHI3L1, EGF, FGF1, GFRA1, GPR56, GPRK6, IL10RB, IL10RA, IL12RB1, KCNJ14,
CC KCNQ1, ORCTL4, PDGFRA, SCYB14, SLC12A1, SLC2A3, TGFBR3, TMEM1, CALCR,
CC IL17R, OSTF1, FGF6, HGF, MET, TGFBI and VEGF, or detecting the base in a
CC polymorphism region existing in the vicinity of any one of the genes. The
CC invention further comprises a reagent or kit for testing secondary
CC hyperparathyroidism in a chronic renal failure patient. This
CC polymucleotide sequence represents the polymorphism containing human
CC IL12RB1 gene of the invention.

XX Sequence 37442 BP; 9374 A; 9810 C; 9241 G; 9017 T; 0 U; 0 Other;

XX Query Match 70.4%; Score 17.6; DB 14; Length 37442;

XX Best Local Similarity 83.3%; Pred. No. 9.5e+02;

XX Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 AAAAACTGGAATCTCAGGCTGAG 24

|||||
Db 18487 AAAAATCTGGAATTACAGGCTGAG 18464

RESULT 31

AEB96540/C

ID AEB96540 standard; DNA; 37442 BP.

XX AEB96540;
AC 06-OCT-2005 (first entry)
XX Human IL12RB1 gene, SEQ ID 24.
XX hepatitis C virus infection; antiinflammatory; hepatotropic; virucide;
KW liver cirrhosis; fibrosis; hepatoma; SNP detection; IL12RB1; ds.
XX Homo sapiens.
XX Key Location/Qualifiers
FH variation 107
FT /*tag= a
FT /standard_name= "Single nucleotide polymorphism"
FT variation 109
FT /*tag= b
FT /standard_name= "Single nucleotide polymorphism"
FT variation 191
FT /*tag= c
FT /standard_name= "Single nucleotide polymorphism"
FT variation 769
FT /*tag= d
FT /standard_name= "Single nucleotide polymorphism"
FT variation 1160
FT /*tag= e
FT /standard_name= "Single nucleotide polymorphism"
FT variation 1693
FT /*tag= f
FT /standard_name= "Single nucleotide polymorphism"
FT variation 3410
FT /*tag= g
FT /standard_name= "Single nucleotide polymorphism"
FT variation 5893
FT /*tag= h
FT /standard_name= "Single nucleotide polymorphism"
FT variation 6662
FT /*tag= i
FT /standard_name= "Single nucleotide polymorphism"
FT variation 6929
FT /*tag= j
FT /standard_name= "Single nucleotide polymorphism"
FT variation 6989
FT /*tag= k
FT /standard_name= "Single nucleotide polymorphism"
FT variation 7715
FT /*tag= l
FT /standard_name= "Single nucleotide polymorphism"
FT variation 8922
FT /*tag= m
FT /standard_name= "Single nucleotide polymorphism"
FT variation 9091
FT /*tag= n
FT /standard_name= "Single nucleotide polymorphism"
FT variation 9879
FT /*tag= o
FT /standard_name= "Single nucleotide polymorphism"
FT variation 10001
FT /*tag= p
FT /standard_name= "Single nucleotide polymorphism"
FT variation 10043
FT /*tag= q
FT /standard_name= "Single nucleotide polymorphism"
FT variation 10072
FT /*tag= r
FT /standard_name= "Single nucleotide polymorphism"
FT variation 10220
FT /*tag= s
FT /standard_name= "Single nucleotide polymorphism"
FT variation 10258
FT /*tag= t
FT /standard_name= "Single nucleotide polymorphism"

CC The nucleic acid molecules, proteins and antibodies can be used in
CC diagnostic or pharmaceutical compounds e.g. for gene therapy. The methods
CC and nucleic acid molecules, etc. are useful for production of transgenic
CC animals, especially a transgenic mouse for the study of diseases
CC associated with bone development, e.g. spondyloepiphyseal dysplasia and
CC achondroplasia. This sequence encodes the murine LOBO protein described
CC in the method of the invention

SQ Sequence 49999 BP; 13210 A; 11814 C; 10825 G; 14150 T; 0 U; 0 Other;

Query Match 70.4%; Score 17.6; DB 2; Length 49999;
Best Local Similarity 83.3%; Pred. No. 9.9e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AAAAACTGGAATCTCAGGCTGAGA 25
|||||
DB 45957 AAAAACTGGAATCTCAGGCTGAGA 45934

RESULT 35
ADQ97081/C
ID ADQ97081 standard; DNA; 63411 BP.

XX AC ADQ97081;

XX DT 07-OCT-2004 (first entry)

XX DE Mouse cancer associated sequence MD10-004, SEQ ID 57.

XX KW Cytostatic; Gene Therapy; cancer; leukemia; lymphoma; Mouse; ds.

XX OS Mus musculus.

XX PN WO2004060304-A2.

XX PD 22-JUL-2004.

XX PF 22-DEC-2003; 2003WO-US041389.

XX PR 27-DEC-2002; 2002US-00330773.

XX PA (SAGR-) SAGRES DISCOVERY INC.

XX PI Morris DW, Malandro MS;

XX DR WPI; 2004-543781/52.

XX PT New isolated cancer associated nucleic acids comprising at least 10
PT contiguous nucleotides, useful for diagnosing, preventing and/or treating
PT cancers such as leukemia and lymphoma.

XX PS Claim 1; SEQ ID NO 57; 199pp; English.

XX CC The present invention relates to cancer associated sequences (ADQ97025-
CC ADQ98004). The sequences are useful for the diagnosis, prevention and/or
CC treatment of cancer, such as leukemia and lymphoma. Note: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 63411 BP; 17408 A; 15472 C; 14344 G; 15847 T; 0 U; 340 Other;

Query Match 70.4%; Score 17.6; DB 12; Length 63411;
Best Local Similarity 83.3%; Pred. No. 1e+03;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAACTGGAATCTCAGGCTGAG 24
|||||
DB 41743 AAAAACTGGAATCTCAGGCTGAG 41720

RESULT 36
ADC85257

ID ADC85257 standard; DNA; 96594 BP.

XX AC ADC85257;

XX DT 01-JAN-2004 (first entry)

XX DE Mouse Ptpkr genomic sequence.

XX KW Cytostatic; gene therapy; vaccine; cancer; carcinoma-associated gene; CA;
KW secreted; transmembrane; intracellular; ds.

XX OS Mus sp.

XX PN WO2003045230-A2.

XX PD 05-JUN-2003.

XX PF 02-DEC-2002; 2002WO-US038582.

XX PR 30-NOV-2001; 2001US-00997722.

XX PA (SAGR-) SAGRES DISCOVERY.

XX PI Morris DW, Engelhard EK;

XX DR WPI; 2003-513603/48.

XX PT New recombinant nucleic acid comprising a nucleotide sequence of any of
PT the carcinoma-associated (CA) genes, useful for screening for drug
PT candidates for diagnosing or treating carcinomas.

XX PS Claim 1; SEQ ID NO 43; 983pp; English.

XX CC The invention relates to a recombinant nucleic acid comprising a
CC nucleotide sequence selected from any of the fully defined carcinoma-
CC associated (CA) genes from the 50 tables given in the specification. The
CC CA proteins are secreted, transmembrane or intracellular proteins. The
CC recombinant nucleic acids are useful for screening for drug candidates
CC for diagnosing or treating carcinomas. Sequences given in ADC85215-
CC ADC85514 represent CA genes of the invention.

XX SQ Sequence 96594 BP; 27993 A; 18068 C; 18732 G; 30838 T; 0 U; 963 Other;

Query Match 70.4%; Score 17.6; DB 10; Length 96594;
Best Local Similarity 83.3%; Pred. No. 1.1e+03;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAACTGGAATCTCAGGCTGAG 24

DB 27319 AATAATCTGTAATCTCAGGCTGAG 27342

RESULT 37

ADA02777

ID ADA02777 standard; DNA; 96595 BP.

XX AC ADA02777;

XX DT 06-NOV-2003 (first entry)

XX DE Mouse Ptpkr carcinoma associated gene, SEQ ID NO:1295.

XX KW Mouse; murine; carcinoma associated; oncogene; carcinoma; cancer; breast;
KW prostate; lymphoma; leukaemia; cytostatic; gene therapy; drug screening;
KW gene; ds.

XX OS Mus sp.

XX PN WO2003057146-A2.

XX PD 17-JUL-2003.

XX PF 26-DEC-2002; 2002WO-US041414.

```

XX PR 26-DEC-2001; 2001US-00035832.
XX PA (SAGR-) SAGRES DISCOVERY.
XX PI Morris DW;
XX WPI; 2003-587068/55.
XX PT New recombinant nucleic acid encoding carcinoma associated protein,
XX useful for preparing compositions for treating carcinomas.
XX PS Claim 1; SEQ ID NO 1295; 245pp; English.
XX CC The invention relates to recombinant carcinoma associated (CA) nucleic
XX acid sequences from mouse and human (ADA01482-ADA03094), and to
XX recombinant carcinoma associated proteins (CAP) encoded by them. The
XX invention also encompasses expression vectors and host cells comprising a
XX CA nucleic acid, a polypeptide (especially an antibody) that specifically
XX binds to the protein, and a biochip comprising CA nucleic acid or
XX fragments thereof. The sequences of the invention were identified using
XX oncogenic retroviruses, which insert into the genome of the host organism
XX at random. Many of these do not carry transduced host oncogenes or
XX pathogenic trans-acting viral genes, meaning that cancer incidence is a
XX direct consequence of the effects of proviral integration into host
XX protooncogenes. The CA nucleic acid sequences can be used to diagnose
XX carcinoma (especially breast cancer, prostate cancer, lymphoma or
XX leukaemia) or a propensity to carcinoma by determination of the sequence
XX of a CA gene, or by determination of CA gene expression in particular
XX tissues. CA nucleic acids, proteins and antibodies are also useful as
XX therapeutic agents and in screening and evaluating drug candidates. The
XX present sequence represents a specifically claimed murine CA nucleic acid
XX sequence of the invention. Note: The complete sequence data for this
XX patent did not form part of the printed specification, but was obtained
XX in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 96595 BP; 27994 A; 18068 C; 18732 G; 30838 T; 0 U; 963 Other;

Query Match 70.4%; Score 17.6; DB 9; Length 96595;
Best Local Similarity 83.3%; Pred. No. 1.1e+03;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAACTGGAATCTCAGGCTGAG 24
DB 27319 ATAAATCTGTAATCTCAGGCTGAG 27342

RESULT 38
ADB72515
ID ADB72515 standard; DNA; 96595 BP.
XX AC ADB72515;
XX DT 04-DEC-2003 (first entry)
XX DE Mouse Ptpk gene.
XX KW mouse; ds; cytostatic; gene therapy; vaccine; carcinoma; lymphomas;
XX KW cancer; neoplasm; adenocarcinoma; sarcoma; gene.
XX OS Mus sp.
XX PN WO2003008583-A2.
XX PD 30-JAN-2003.
XX PF 26-DEC-2001; 2001WO-US051291.
XX PR 02-OCT-2001; 2001US-00798586.
XX PR 23-OCT-2001; 2001US-00004113.
XX PR 08-NOV-2001; 2001US-00052482.
XX PR 30-NOV-2001; 2001US-00997722.

```

```

PR 20-DEC-2001; 2001US-00034650.
XX (SAGR-) SAGRES DISCOVERY.
XX PA Morris DW, Engelhard EK;
XX PI Morris DW;
XX WPI; 2003-239337/23.
XX PT New recombinant nucleic acid, useful for treating carcinomas, lymphomas,
XX cancers, neoplasm, adenocarcinoma, or sarcomas.
XX PS Claim 1; SEQ ID NO 343; 2304pp; English.
XX CC The invention relates to a novel recombinant nucleic acid comprising a
XX nucleotide sequence selected from any of the 660 sequences fully defined
XX in the specification. A polynucleotide of the invention has cytostatic
XX activity, and may have a use in gene therapy, or in a vaccine. The
XX recombinant nucleic acids and polypeptides are useful for treating
XX carcinomas, e.g. lymphomas, cancers, neoplasm, adenocarcinoma, and
XX sarcomas. The present sequence represents a mouse gene of the invention.
XX SQ Sequence 96595 BP; 27994 A; 18068 C; 18732 G; 30838 T; 0 U; 963 Other;

Query Match 70.4%; Score 17.6; DB 10; Length 96595;
Best Local Similarity 83.3%; Pred. No. 1.1e+03;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAACTGGAATCTCAGGCTGAG 24
DB 27319 ATAAATCTGTAATCTCAGGCTGAG 27342

RESULT 39
ADM74372
ID ADM74372 standard; DNA; 96595 BP.
XX AC ADM74372;
XX DT 01-JUL-2004 (first entry)
XX DE Murine carcinoma associated (CA) nucleic acid #22.
XX KW Mouse; carcinoma associated nucleic acid; CA nucleic acid; gene; ds;
XX KW carcinoma associated protein; CAP; carcinoma; leukaemia; lymphoma;
XX KW cytostatic.
XX OS Mus musculus.
XX PN US2004072154-A1.
XX PD 15-APR-2004.
XX PF 30-NOV-2001; 2001US-00997722.
XX PR 22-DEC-2000; 2000US-00747377.
XX PR 02-MAR-2001; 2001US-00798586.
XX PA (MORR/) MORRIS D W.
XX PA (ENGE/) ENGELHARD E K.
XX PI Morris DW, Engelhard EK;
XX WPI; 2004-328562/30.
XX PT New carcinoma associated gene or protein, useful for preparing a
XX composition for diagnosing or treating carcinoma e.g., leukemia or
XX lymphoma.
XX PS Claim 1; SEQ ID NO 43; 29pp; English.
XX CC The invention relates to new recombinant nucleic acids. The invention
XX also relates to a host cell comprising a recombinant nucleic acid or
XX expression vector, an expression vector comprising a recombinant nucleic

```

CC acid, a recombinant protein, a method of screening for drug candidates, a
CC method of screening for a bioactive agent capable of binding to a
CC carcinoma associated protein (CAP) encoded by a nucleotide sequence, a
CC method of screening for a bioactive agent capable of modulating the
CC activity of a CAP, a method of evaluating the effect of a candidate
CC carcinoma drug, a method of diagnosing carcinoma, a method for inhibiting
CC the activity of a CAP, a method of treating carcinomas, a method of
CC neutralising the effect of a CAP and a method of diagnosing carcinoma or
CC propensity to carcinoma. A method of evaluating the effect of a candidate
CC carcinoma drug comprises administering the drug to a patient, removing a
CC cell sample from the patient and determining alterations in the
CC expression or activation of a gene comprising the nucleotide sequence. A
CC method of diagnosing carcinoma comprises determining the expression of
CC one or more genes comprising the nucleic acid sequence in a first tissue
CC type of a first individual and comparing the expression of the gene from
CC a second normal tissue type from the first individual or a second
CC unaffected individual, where a difference in the expression indicates
CC that the first individual has carcinoma. A method of inhibiting the
CC activity of a CAP comprises binding an inhibitor to the CAP. Treating
CC carcinomas comprises administering to a patient an inhibitor of CAP.
CC Neutralising the effect of a CAP comprises contacting an agent specific
CC for the CAP. The polypeptide specifically binds to the protein encoded by
CC the nucleic acid. It comprises an antibody that specifically binds to the
CC protein encoded by the nucleic acid. The nucleic acids are useful for
CC preparing a composition for diagnosing or treating carcinoma e.g.,
CC leukaemia or lymphoma. This sequence represents a murine carcinoma
CC associated (CA) nucleic acid of the invention. Note: The sequence data
CC for this patent did not form part of the printed specification but was
CC obtained in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html.

XX SQ Sequence 96595 BP; 27994 A; 18068 C; 18732 G; 30838 T; 0 U; 963 Other;

Query Match 70.4%; Score 17.6; DB 12; Length 96595;

Best Local Similarity 83.3%; Pred. No. 1.1e+03;

Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAACTGGAATCTCAGGCTGAG 24

Db 27319 ATAAATCTGTAATCTCAGGCTGAG 27342

RESULT 40

ACN43998_2/C
Continuation (3 of 7) of ACN43998 from base 200001 (Human genomic sequence HCG1778483.)

WP Sequence split into 7 fragments LOCUS ACN43998 Accession ACN43998

WP	Fragment Name	Begin	End
WP	ACN43998_0	1	110000
WP	ACN43998_1	100001	210000
WP	ACN43998_2	200001	310000
WP	ACN43998_3	300001	410000
WP	ACN43998_4	400001	510000
WP	ACN43998_5	500001	610000
WP	ACN43998_6	600001	653122

Query Match 70.4%; Score 17.6; DB 11; Length 110000;

Best Local Similarity 83.3%; Pred. No. 1.1e+03;

Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AAAAACTGGAATCTCAGGCTGAGA 25

Db 51517 AAAAATAAACATCTCAGACTGAGA 51494

RESULT 41

ACN45090_0

WP Sequence split into 4 fragments LOCUS ACN45090 Accession ACN45090

WP	Fragment Name	Begin	End
WP	ACN45090_0	1	110000
WP	ACN45090_1	100001	210000
WP	ACN45090_2	200001	310000
WP	ACN45090_3	300001	350764

ID ACN45090 standard; DNA; 350764 BP.

XX ACN45090;
AC 18-NOV-2004 (first entry)
DT Human genomic sequence HCG22125.
DE Cytostatic; carcinoma; lymphoma; cancer; human; gene; ss.
KW Homo sapiens.
OS WO2003073826-A2.
PN 12-SEP-2003.
PD 28-FEB-2003; 2003WO-US006235.
PF 01-MAR-2002; 2002US-00087192.
PR (SAGR-) SAGRES DISCOVERY.
PA Morris DW;
PI WPI; 2003-328604/31.
DR Recombinant nucleic acid useful for diagnosis and treatment of carcinoma
PT comprises a nucleotide sequence.
PS Claim 1; SEQ ID NO 1864; Opp; English.

XX The present invention relates to novel DNA and protein sequences which

are associated with carcinomas. The sequences are useful for: (i) for
screening drug candidates; (ii) for screening of bioactive agent capable
of binding to Carcinoma Associated Protein (CAP); (iii) for screening of
a bioactive agent capable of modulating the activity of CAP; (iv) for
evaluating the effect of a candidate carcinoma drug; (v) for diagnosing
carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating
carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip;
(x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for
determining Carcinoma Associated (CA) gene copy number. In addition, the
CA genes are useful as DNA vaccines and the CAP are useful as markers of
carcinoma including lymphoma. The present sequence is one such CA coding
sequence. Note: This patent is an equivalent to basic patent
US2002182586A1, for which no sequence data was published

XX SQ Sequence 350764 BP; 99569 A; 63584 C; 68015 G; 119596 T; 0 U; 0 Other;

Query Match 70.4%; Score 17.6; DB 11; Length 110000;

Best Local Similarity 83.3%; Pred. No. 1.1e+03;

Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAACTGGAATCTCAGGCTGAG 24

Db 60025 AAAATGCTGCATCTCTCAGGCTGAG 60048

RESULT 42

ABQ88198

ID ABQ88198 standard; cDNA; 154902 BP.

XX ABQ88198;

XX 18-SEP-2002 (first entry)

XX Human osteoblast differentiation related cDNA SEQ ID NO 105.

XX Human; osteoblast; stem cell differentiation; bone tissue deposition;

XX osteoporosis; osteopathic; ss.

XX Homo sapiens.

XX WO200250301-A2.

PD 27-JUN-2002.
XX
XX
PF 18-DEC-2001; 2001WO-US048276.
XX
XX
PR 18-DEC-2000; 2000US-0255882P.
PR 24-APR-2001; 2001US-0285691P.
XX
XX
PA (GENE-) GENE LOGIC INC.
XX (PROC) PROCTER & GAMBLE CO.
XX
XX
PI Ji D, Axelrod DW, Cook JS, Jaiswal N, Einstein R, Houghton A;
PI Mertz L;
XX
XX
DR WPI; 2002-557663/59.
XX
XX
PT Use of genes and their expression profiles associated with osteoblast
PT differentiation for screening modulators bone formation, for diagnosing
PT or treating e.g. osteoporosis, or as markers for the differentiation
XX process.
XX
PS Claim 1; SEQ ID NO 105; 78pp + Sequence Listing; English.
XX
XX
CC The invention relates to genes and their expression profiles are used
CC for: (a) screening modulators of precursor stem cell differentiation into
CC osteoblasts, or bone tissue deposition; (b) diagnosing abnormal
CC deposition of bone tissue, abnormal rate of osteoblast formation or
CC osteoporosis; or (c) treating or monitoring treatment of the conditions
CC cited in (b), or monitoring the progression of bone tissue deposition.
CC Specific conditions include postmenopausal osteoporosis, glucocorticoid
CC osteoporosis or male osteoporosis, osteopenia, osteodystrophy, drug-
CC induced abnormalities in bone formation or bone loss, conditions that
CC involve altered bone metabolism (e.g. idiopathic juvenile osteoporosis),
CC skeletal disease linked to breast cancer, mastocytosis, Fanconi syndrome
CC or fibrous dysplasia. The present sequence is that of an osteoblast
CC differentiation associated cDNA marker of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
XX
SQ Sequence 154902 BP; 43917 A; 31458 C; 32848 G; 46679 T; 0 U; 0 Other;
Query Match 70.4%; Score 17.6; DB 6; Length 154902;
Best Local Similarity 83.3%; Pred. No. 1.2e+03;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 2 AAAAACTGGAATCTCAGGCTGAGA 25
DB 33872 ATAATCTGGATCTCAGGCTGAGA 33895
RESULT 43
ADZ12614
ID ADZ12614 standard; DNA; 289106 BP.
XX
XX
AC ADZ12614;
XX
XX
DT 16-JUN-2005 (first entry)
XX
XX
DE Murine cancer-associated genomic DNA #13.
XX
XX
KW Diagnosis; DNA microarray; microarray; biochip; cancer; neoplasm;
KW Cytostatic; Gene; ds.
XX
XX
OS Mus sp.
XX
XX
PN WO2005031001-A2.
XX
XX
PD 07-APR-2005.
XX
XX
PF 23-SEP-2004; 2004WO-US031617.
XX
XX
PR 23-SEP-2003; 2003US-00669920.
XX

PA (CHTR) CHIRON CORP.
XX
XX
PI Morris DW, Malandro MS;
XX
XX
DR WPI; 2005-273395/28.
XX
XX
PT Nucleic acid array useful for detecting cancer associated nucleic acid,
PT comprises two or more nucleic acid probes.
XX
XX
PS Disclosure; SEQ ID NO 134; 198pp; English.
XX
XX
CC The invention relates to a nucleic acid array for detecting a cancer
CC associated (CA) nucleic acid, comprising two or more nucleic acid probes.
CC The invention also relates to a peptide array comprising two or more
CC isolated polypeptides encoded by a CA nucleic acid sequence, a compound
CC that binds to a polypeptide, an isolated antibody or its fragment which
CC binds to a polypeptide, which is prepared by immunizing a host animal
CC with a composition comprising the polypeptide or its antigen binding
CC fragment and collecting cells from the host expressing antibodies against
CC the antigen or its antigen binding fragment, a composition comprising the
CC antibody and a carrier, a method of screening for anticancer activity, a
CC method of detecting a CA nucleic acid, a method of diagnosing cancer, a
CC method of treating cancer and a method of inhibiting expression of a CA
CC nucleic acid in a cell. The CA nucleic acids are useful for detecting CA
CC nucleic acids. The antibody is useful for detecting the presence or
CC absence of cancer cells in an individual which involves contacting cells
CC from the individual with the antibody and detecting a complex of a CA
CC protein from the cancer cells and the antibody, where the detection of
CC the complex correlates with the presence of cancer cells in the
CC individual. The composition is useful for inhibiting growth of cancer
CC cells in an individual or for delivering a therapeutic agent to cancer
CC cells in an individual. The invention is also useful for diagnosing
CC cancer, for treating cancer and for inhibiting expression of a CA gene in
CC a cell. This sequence represents murine cancer-associated genomic DNA of
CC the invention.
XX
XX
SQ Sequence 289106 BP; 84569 A; 54316 C; 55029 G; 92921 T; 0 U; 2271 Other;
Query Match 70.4%; Score 17.6; DB 14; Length 289106;
Best Local Similarity 83.3%; Pred. No. 1.3e+03;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 AAAAACTGGAATCTCAGGCTGAG 24
DB 84553 ATAATCTGTAATCTCAGGCTGAG 84576
RESULT 44
ADQ59380/c
ID ADQ59380 standard; DNA; 299598 BP.
XX
XX
AC ADQ59380;
XX
XX
DT 07-OCT-2004 (first entry)
XX
XX
DE Human cancer-associated (CA) gene sequence SEQ ID NO:16.
XX
XX
KW human; cancer-associated gene; cancer-associated protein; cytostatic;
KW gene therapy; vaccine; tyrosine kinase antagonist;
KW G-protein coupled receptor antagonist; cancer; lymphoma; gene; ds.
XX
XX
OS Homo sapiens.
XX
XX
PN WO2004058288-A1.
XX
XX
PD 15-JUL-2004.
XX
XX
PF 15-DEC-2003; 2003WO-US040082.
XX
XX
PR 17-DEC-2002; 2002US-00322696.
XX
XX
PA (SAGR-) SAGRES DISCOVERY INC.
XX

```
PI Morris DW, Malandro MS;
XX
XX WPI: 2004-543349/52.
DR P-PSDB; ADQ59382.
XX
XX New cancer-associated nucleic acid for diagnosing, preventing or treating
PT cancer (e.g. lymphoma) or for screening agents that may be used for
PT treating or preventing cancer.
XX
XX Claim 16; SEQ ID NO 16; 143pp; English.
XX
XX The present invention describes human cancer-associated (CA) nucleotide
CC sequences (1). Also described: (1) an expression vector comprising (1);
CC (2) a host cell comprising (1) or the expression vector; (3) a microarray
CC for detecting a CA nucleic acid; (4) an isolated polypeptide encoded
CC within an open reading frame of a CA sequence; (5) an isolated antibody,
CC or its antigen binding fragment, that binds to the above polypeptide; (6)
CC a hybridoma that produces the monoclonal antibody described above; (7) a
CC pharmaceutical composition comprising the antibody and a pharmaceutical
CC excipient; (8) a kit for detecting or diagnosing cancer cells, comprising
CC the above (monoclonal) antibody or polynucleotide that selectively
CC hybridizes to any of the polynucleotide sequences mentioned above; (9)
CC methods for diagnosing cancer or for detecting the presence or absence of
CC cancer cells in an individual; (10) a method for inhibiting growth of
CC cancer cells in an individual; (11) a method for delivering a therapeutic
CC agent to cancer cells in an individual; (12) an electronic library
CC comprising the polynucleotide or polypeptide, or their fragments,
CC mentioned above; (13) a method of screening for anticancer activity; (14)
CC methods for detecting cancer associated with expression of a polypeptide
CC or the presence of the antibody in a test cell or serum sample; (15) a
CC method for screening for a bioactive agent capable of modulating the
CC activity of a CA protein encoded by the above nucleic acid molecule; and
CC (16) a method for treating cancers. (1) has cytostatic activity, and can
CC be used in gene therapy, in vaccines, as a tyrosine kinase antagonist,
CC and as a G-protein coupled receptor antagonist. The compositions and
CC methods of the present invention can be used for diagnosing, preventing
CC and treating cancer, especially lymphomas. They may also be used in
CC screening for agents that may be used for treating or preventing cancer.
CC The present sequence represents a human CA gene sequence, which is given
CC in the exemplification of the present invention. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 299598 BP; 92218 A; 54563 C; 55801 G; 95585 T; 0 U; 1431 Other;
SQ
Query Match 70.4%; Score 17.6; DB 12; Length 299598;
Best Local Similarity 83.3%; Pred. No. 1.3e+03;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AAAAACTGGAAATCTCAGGCTGAGA 25
| | | | | | | | | | | | | | | | | | | | |
Db 297755 AAAAACTGAAAAATCAGGCTTAGA 297732

RESULT 45
AD081499
ID AD081499 standard; cDNA; 1059 BP.
XX
XX AD081499;
XX
XX 21-APR-2005 (first entry)
XX
XX Plant full length insert polynucleotide seqid 219.
XX
XX plant protectant; plant growth regulant; gene therapy; plant;
XX recombinant DNA construct; physical array; plant breeding marker;
XX cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
XX extreme osmotic condition; pathogen tolerance; pest tolerance;
XX growth rate; cell cycle pathway; disease resistance;
XX galactomannan production; lignin production; plant growth regulator;
XX yield; plant growth; plant development; seed oil; protein yield;
XX protein content; gene; ss.
```

```
XX Unidentified.
OS US2004034888-A1.
PN
XX 19-FEB-2004.
PD
XX 28-APR-2003; 2003US-00425114.
PF
XX 06-MAY-1999; 99US-00304517.
PR 05-NOV-2001; 2001US-00985678.
XX
XX (LIUJ/) LIU J.
PA (ZHOU/) ZHOU Y.
PA (KOVA/) KOVALIC D K.
PA (SCRE/) SCREEN S E.
PA (TABR/) TABASKA J E.
PA (CAOY/) CAO Y.
XX
XX Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
PI WPI: 2004-180133/17.
XX
XX New recombinant DNA construct, useful for improving plant tolerance to
DR cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
XX pests, for conferring increased resistance to plant disease, or for
XX improving yield.
XX
XX Claim 1; SEQ ID NO 219; 15pp; English.
XX
XX The invention describes a recombinant DNA construct comprising a
CC polynucleotide consisting of a sequence encoding an amino acid sequence
CC available in electronic form from the US patent office at
CC ftp.segdata.uspto.gov/sequence.html?DocID:2004034888. The polynucleotide
CC of the invention are also useful in physical arrays of molecules and as
CC plant breeding markers. The recombinant DNA construct is useful for
CC improving plant tolerance to cold, heat, drought, herbicides, extreme
CC osmotic conditions, pathogens or pests, for manipulating growth rate in
CC plant cells by modification of the cell cycle pathway, for conferring
CC increased resistance to plant disease, for producing galactomannan,
CC lignin or plant growth regulators, for increasing the rate of homologous
CC recombination in plants, for improving yield by modification of
CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
CC or by providing improved plant growth and development under at least one
CC stress condition or for modifying seed oil or protein yield and/or
CC content. This sequence represents a plant full length insert
CC polynucleotide that can be used in the recombinant DNA construct of the
CC invention.
XX
XX Sequence 1059 BP; 323 A; 235 C; 266 G; 235 T; 0 U; 0 Other;
SQ
Query Match 69.6%; Score 17.4; DB 13; Length 1059;
Best Local Similarity 94.7%; Pred. No. 6.6e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AAAAACTGGAAATCTCAGGC 20
| | | | | | | | | | | | | | | | | | | | |
Db 49 AAAAGCTGGAATCTCAGGC 67

RESULT 46
AD083966
ID AD083966 standard; cDNA; 2206 BP.
XX
XX AD083966;
XX
XX 21-APR-2005 (first entry)
XX
XX Plant full length insert polynucleotide seqid 2686.
XX
XX plant protectant; plant growth regulant; gene therapy; plant;
XX recombinant DNA construct; physical array; plant breeding marker;
XX cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
XX protein content; gene; ss.
```

KW extreme osmotic condition; pathogen tolerance; pest tolerance;
KW growth rate; cell cycle pathway; disease resistance;
KW Galactomannan production; lignin production; plant growth regulator;
KW yield; plant growth; plant development; seed oil; protein yield;
KW protein content; gene; ss.
XX Unidentified.
XX US2004034888-A1.
XX 19-FEB-2004.
XX 28-APR-2003; 2003US-00425114.
XX 06-MAY-1999; 99US-00304517.
XX 05-NOV-2001; 2001US-00985678.
XX (LIUJ/) LIU J.
XX (ZHOU/) ZHOU Y.
XX (KOVA/) KOVALIC D K.
XX (SCRE/) SCREEN S E.
XX (TABA/) TABASKA J E.
XX (CAOY/) CAO Y.
XX Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
XX WPI; 2004-180133/17.
XX New recombinant DNA construct, useful for improving plant tolerance to
XX cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
XX pests, for conferring increased resistance to plant disease, or for
XX improving yield.
XX Claim 1; SEQ ID NO 2686; 15pp; English.
XX The invention describes a recombinant DNA construct comprising a
XX polynucleotide consisting of a sequence encoding an amino acid sequence
XX available in electronic form from the US patent office at
XX ftp.segdata.uspto.gov/sequence.html?docid:2004034888. The polynucleotide
XX of the invention are also useful in physical arrays of molecules and as
XX plant breeding markers. The recombinant DNA construct is useful for
XX improving plant tolerance to cold, heat, drought, herbicides, extreme
XX osmotic conditions, pathogens or pests, for manipulating growth rate in
XX plant cells by modification of the cell cycle pathway, for conferring
XX increased resistance to plant disease, for producing galactomannan,
XX lignin or plant growth regulators, for increasing the rate of homologous
XX recombination in plants, for improving yield by modification of
XX photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
XX or by providing improved plant growth and development under at least one
XX stress condition or for modifying seed oil or protein yield and/or
XX content. This sequence represents a plant full length insert
XX polynucleotide that can be used in the recombinant DNA construct of the
XX invention.
XX Sequence 2206 BP; 691 A; 436 C; 547 G; 532 T; 0 U; 0 Other;
Query Match 69.6%; Score 17.4; DB 13; Length 2206;
Best Local Similarity 94.7%; Pred. No. 7.5e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 AAAAAGCTGGAATCTCAGGC 20
|||||
Db 1215 AAAAAGCTGGAATCTCAGGC 1233
RESULT 47
ADX63913
ID ADX63913 standard; cDNA; 2800 BP.
XX
XX AC ADX63913;
XX
XX DT 21-APR-2005 (first entry)
XX

DE Plant full length insert polynucleotide seqid 34756.
XX plant protectant; plant growth regulant; gene therapy; plant;
XX recombinant DNA construct; physical array; plant breeding marker;
KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
KW extreme osmotic condition; pathogen tolerance; pest tolerance;
KW growth rate; cell cycle pathway; disease resistance;
KW galactomannan production; lignin production; plant growth regulator;
KW yield; plant growth; plant development; seed oil; protein yield;
KW protein content; gene; ss.
XX Unidentified.
XX US2004034888-A1.
XX 19-FEB-2004.
XX 28-APR-2003; 2003US-00425114.
XX 06-MAY-1999; 99US-00304517.
XX 05-NOV-2001; 2001US-00985678.
XX (LIUJ/) LIU J.
XX (ZHOU/) ZHOU Y.
XX (KOVA/) KOVALIC D K.
XX (SCRE/) SCREEN S E.
XX (TABA/) TABASKA J E.
XX (CAOY/) CAO Y.
XX Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
XX WPI; 2004-180133/17.
XX New recombinant DNA construct, useful for improving plant tolerance to
XX cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
XX pests, for conferring increased resistance to plant disease, or for
XX improving yield.
XX Claim 1; SEQ ID NO 34756; 15pp; English.
XX The invention describes a recombinant DNA construct comprising a
XX polynucleotide consisting of a sequence encoding an amino acid sequence
XX available in electronic form from the US patent office at
XX ftp.segdata.uspto.gov/sequence.html?docid:2004034888. The polynucleotide
XX of the invention are also useful in physical arrays of molecules and as
XX plant breeding markers. The recombinant DNA construct is useful for
XX improving plant tolerance to cold, heat, drought, herbicides, extreme
XX osmotic conditions, pathogens or pests, for manipulating growth rate in
XX plant cells by modification of the cell cycle pathway, for conferring
XX increased resistance to plant disease, for producing galactomannan,
XX lignin or plant growth regulators, for increasing the rate of homologous
XX recombination in plants, for improving yield by modification of
XX photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
XX or by providing improved plant growth and development under at least one
XX stress condition or for modifying seed oil or protein yield and/or
XX content. This sequence represents a plant full length insert
XX polynucleotide that can be used in the recombinant DNA construct of the
XX invention.
XX Sequence 2800 BP; 835 A; 622 C; 692 G; 651 T; 0 U; 0 Other;
Query Match 69.6%; Score 17.4; DB 13; Length 2800;
Best Local Similarity 94.7%; Pred. No. 7.8e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 AAAAAGCTGGAATCTCAGGC 20
|||||
Db 1801 AAAAAGCTGGAATCTCAGGC 1819
RESULT 48
ABA20105/c
ID ABA20105 standard; DNA; 5161 BP.

XX ABA20105;
XX AC
XX DT
XX 23-JAN-2002 (first entry)
XX DE
XX DE Human nervous system related polynucleotide SEQ ID NO 12436.
XX DE
XX DE Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;
KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;
KW antiparkinsonian; antispasmodic; antianemic; antiarthritic; cancer;
KW antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
KW antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;
KW antiparasitic; cardiant; immune disorder; cardiovascular disorder;
KW neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.
XX
OS Homo sapiens.
XX
XX WO200159063-A2.
XX
XX 16-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US001334.
XX
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184564P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 06-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 20-OCT-2000; 2000US-0242221P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246521P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246603P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.

PR 01-DEC-2000; 2000US-0250391P.
PR 01-DEC-2000; 2000US-0251160P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251990P.
PR 08-DEC-2000; 2000US-0251999P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-541565/60.
XX
XX Nucleic acids encoding 3224 human nervous system antigen polypeptides,
PT useful for preventing, diagnosing and/or treating nervous system cancers
PT and metastases.
XX
XX Disclosure; SEQ ID NO 12436; 1701pp + Sequence Listing; English.
XX
XX The invention relates to novel genes (ABAI1004-ABA21534) and proteins
CC (ABBI4678-ABBI8001) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in
CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
CC cardiovascular disorders such as myocardial ischaemias; (d) wound healing
CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
CC infectious diseases such as viral, bacterial, fungal and parasitic
CC infections. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX
XX Sequence 5161 BP; 1167 A; 1428 C; 1321 G; 1245 T; 0 U; 0 Other;
Query Match 69.6%; Score 17.4; DB 5; Length 5161;
Best Local Similarity 94.7%; Pred. No. 8.5e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 7 CTGGAATCTCAGGCTGAGA 25
Db 2183 CTGGAATCTCAGGCTGAGA 2165
RESULT 49
ACN43916/c
ID ACN43916 standard; DNA; 35456 BP.
XX
XX ACN43916;
XX
XX 18-NOV-2004 (first entry)
XX
XX Mouse genomic sequence mCG11078.
XX
XX Cytostatic; carcinoma; lymphoma; cancer; murine; gene; ss.
XX
XX Mus musculus.
XX
XX WO2003073826-A2.
XX
XX 12-SEP-2003.
XX

PF 28-FEB-2003; 2003WO-US006235.
XX
PR 01-MAR-2002; 2002US-00087192.
XX
XX (SAGR-) SAGRES DISCOVERY.
XX
XX Morris DW;
XX
XX WPI; 2003-328604/31.
XX
XX Recombinant nucleic acid useful for diagnosis and treatment of carcinoma
PT comprises a nucleotide sequence.
XX
XX Claim 1; SEQ ID NO 103; Opp; English.
XX
XX The present invention relates to novel DNA and protein sequences which
CC are associated with carcinomas. The sequences are useful for: (i) for
CC screening drug candidates; (ii) for screening of bioactive agent capable
CC of binding to Carcinoma Associated Protein (CAP); (iii) for screening of
CC a bioactive agent capable of modulating the activity of CAP; (iv) for
CC evaluating the effect of a candidate carcinoma drug; (v) for diagnosing
CC carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating
CC carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip;
CC (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for
CC determining Carcinoma Associated (CA) gene copy number. In addition, the
CC CA genes are useful as DNA vaccines and the CAP are useful as markers of
CC carcinoma including lymphoma. The present sequence is one such CA coding
CC sequence. Note: This patent is an equivalent to basic patent
CC US2002182586A1, for which no sequence data was published
XX
XX Sequence 35456 BP; 7528 A; 8218 C; 7732 G; 8394 T; 0 U; 3584 Other;
SQ
Query Match 69.6%; Score 17.4; DB 11; Length 35456;
Best Local Similarity 94.7%; Pred. No. 1.2e+03;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 5 AACTGGAATCTCAGGCTGA 23
Db 35037 AACTGGAATCTCAGGCTGA 35019
RESULT 50
ADP94079
ID ADP94079 standard; cDNA; 460 BP.
XX
XX ACN43916;
XX
XX 09-SEP-2004 (first entry)
XX
XX Cotton expressed sequence tag, EST, #3090.
XX
XX Cotton; ss; EST; expressed sequence tag; plant; plant protection;
XX plant improvement; marker-assisted breeding.
XX
XX Gossypium hirsutum; variety Nucleon33B.
XX
XX US2004123338-A1.
XX
XX 24-JUN-2004.
XX
XX 08-DEC-2000; 2000US-00732627.
XX
XX 10-DEC-1999; 99US-0170255P.
XX
XX (FINC/) FINCHER K L.
XX
XX Fincher KL;
XX
XX WPI; 2004-479807/45.
XX
XX New substantially purified nucleic acid molecule that encodes a cotton
XX protein or its fragment, useful as molecular tool for the targeting and
XX isolation of novel genes for plant protection and improvement.
XX

XX Claim 1; SEQ ID NO 3090; 30pp; English.
PS
XX
CC The invention relates to a substantially purified nucleic acid molecule
CC that encodes a cotton protein or its fragment comprising an EST
CC (expressed sequence tag) appearing as ADP9090-ADP95919. Also included
CC are a substantially purified cotton protein or its fragment encoded by a
CC nucleic acid molecule above and a transformed plant (having a nucleic
CC acid molecule which comprises: an exogenous promoter region which
CC functions in a plant cell to cause the production of a mRNA molecule; a
CC structural nucleic acid molecule comprising one of the ESTs or their
CC complements; a 3' non-translated sequence that functions in the plant
CC cell to cause termination of transcription and addition of polyadenylated
CC ribonucleotides to a 3' end of the mRNA molecule). The ESTs are useful as
CC molecular tool for the targeting and isolation of novel genes for plant
CC protection and improvement. The ESTs are useful for developing new
CC strategies for understanding critical plant developmental and metabolic
CC pathways, for isolating genes and promoters, for identifying and mapping
CC the genes involved in developmental and metabolic pathways, and for
CC determining gene function. The cotton nucleic acid molecules are useful
CC as molecular tags to isolate genetic regions, isolate genes, map genes,
CC and determine gene function. The nucleic acid molecules are useful for
CC determining if genes are members of a particular gene family and for use
CC in marker-assisted breeding programs. The present sequence is one of the
CC 4930 cotton ESTs of the invention. NOTE: The sequences are not displayed
CC in the specification but are available in electronic format from the
CC USPTO at seqdata.uspto.gov/sequence.html?docID=20040123338.
XX
SQ Sequence 460 BP; 127 A; 82 C; 110 G; 141 T; 0 U; 0 Other;

Query Match 68.8%; Score 17.2; DB 12; Length 460;
Best Local Similarity 86.4%; Pred. No. 7.2e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 AAAAAAAGTGGATCTCAGGCTG 22
Db 79 AAAAAAAGTGGATCTCAGGCTG 100

Search completed: February 3, 2006, 21:57:13
Job time : 211.111 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 3, 2006, 14:38:25 ; Search time 2951.67 Seconds
(without alignments)
396.277 Million cell updates/sec

Title: US-10-719-900-30
Perfect score: 25
Sequence: 1 aaaaactggaatctcaggctgaga 25

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 150 summaries

Database :

- EST:*
- 1: gb_est1:*
 - 2: gb_est2:*
 - 3: gb_est3:*
 - 4: gb_hic:*
 - 5: gb_est4:*
 - 6: gb_est5:*
 - 7: gb_est6:*
 - 8: gb_est7:*
 - 9: gb_gss1:*
 - 10: gb_gss2:*
 - 11: gb_gss3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	25	100.0	391	5	BX632940
C 2	25	100.0	435	1	AA963991
C 3	25	100.0	494	2	BE863633
C 4	25	100.0	587	7	CN704489
C 5	25	100.0	771	7	CN461714
C 6	25	100.0	808	2	BE268998
C 7	25	100.0	1717	4	AK043826
C 8	24	96.0	395	5	BY618046
C 9	23.4	93.6	693	10	AG325768
C 10	21.8	87.2	291	2	BB347802
C 11	21.4	85.6	304	2	BB391308
C 12	20.8	83.2	632	6	CA233715
C 13	20.8	83.2	633	6	CA220753
C 14	20.8	83.2	649	6	CA206779
C 15	20.2	80.8	1069	2	BG572464
C 16	20	80.0	362	2	BF589034
C 17	20	80.0	460	2	BG013717
C 18	20	80.0	519	1	AI473209
C 19	20	80.0	581	3	BP250892
C 20	20	80.0	767	5	BX951384
C 21	19.8	79.2	661	9	AQ378916
C 22	19.8	79.2	685	3	BP738553

Consortium at LLNL (info@image.llnl.gov). IMAGE ID=1780854

Seq primer: M13 Forward

FEATURES

Location/Qualifiers

```
1. 435
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-C0-98-9-01-0-UI"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="UI-R-C0"
/notes="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; The UI-R-C0
library is a subtracted library derived from the UI-R-A1
and UI-R-E1 libraries. The UI-R-A1 library consisted of a
mixture of individually tagged normalized libraries
constructed from rat placenta, adult lung, brain, liver,
kidney, heart, spleen, ovary, and muscle. The UI-R-E1
library consisted of a mixture of individually tagged
normalized libraries constructed from 8, 12 and 18-day
embryo. The tag is a string of 3-5 nucleotides present
between the Not I site and the oligo-dT track which
allows identification of the library of origin of a clone
within the mixture. The subtracted library (UI-R-C0) was
constructed as follows: PCR amplified cDNA inserts from a
pool of UI-R-A1 and UI-R-E1 clones from which 3' ESTs had
been derived was used as a driver in a hybridization with
the pooled UI-R-A1 and UI-R-E1 library in the form of
single-stranded circles. The remaining single-stranded
circles (subtracted library) was purified by
hydroxyapatite column chromatography, converted to
double-stranded circles and electroporated into DH10B
bacteria (Life Technologies) to generate the UI-R-C0
library. This procedure has been previously described
(Bonaldo, Lennon and Soares, Genome Research 6: 791-806,
1996)"
```

FEATURES

source

Location/Qualifiers

```
1. 494
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CS7BL/6J"
/db_xref="taxon:10090"
/clone="UI-M-BH0-ake-c-06-0-UI"
/dev_stage="27-32 days"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NIH BMAP_M_S1"
/notes="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; The
NIH BMAP_M_S1 library is a subtracted library derived from
a mixture of normalized libraries from ten regions of the
mouse brain (cerebellum, brain stems, olfactory bulbs,
hypothalamus, cortex, amygdala, basal ganglia, pineal
gland, striatum, hippocampus). The driver used for
subtraction consisted of a pool of 20,000 cDNA clones
obtained from non-normalized and normalized libraries of
these ten regions of the mouse brain."
```

ORIGIN

```
Query Match 100.0%; Score 25; DB 2; Length 494;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 1 AAAAACTGGAATCTCAGGCTGAGA 25

Db 296 AAAAACTGGAATCTCAGGCTGAGA 272

ORIGIN

```
Query Match 100.0%; Score 25; DB 1; Length 435;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 1 AAAAACTGGAATCTCAGGCTGAGA 25

Db 314 AAAAACTGGAATCTCAGGCTGAGA 290

RESULT 3

LOCUS

```
BE863633/c
DEFINITION UI-M-BH0-ake-c-06-0-UI.r1 NIH BMAP_M_S1 Mus musculus cDNA clone
UI-M-BH0-ake-c-06-0-UI 5', mRNA sequence.
```

ACCESSION

BE863633

VERSION

BE863633.1 GI:10383871

KEYWORDS

EST.

SOURCE

Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

Sciurognathi; Muroidae; Muridae; Murinae; Mus.

1 (bases 1 to 494)

Bonaldo, M.F., Lennon, G. and Soares, M.B.

Normalization and subtraction: two approaches to facilitate gene

discovery

Genome Res. 6 (9), 791-806 (1996)

Contact: Chin, H

National Institute of Mental Health

6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD

20892-9643, USA

Tel: 301 443 1706

TITLE

JOURNAL

COMMENT

Transcriptome analysis of mouse stem cells and early embryos

PLoS Biol. 1 (3), 410-419 (2003)

Contact: Dawood B. Dudekula

Laboratory of Genetics

National Institute on Aging/National Institutes of Health

333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA

Email: cdna@igsun.grc.nia.nih.gov

Plate: E0488 row: D column: 11

Seq primer: M13 Reverse

High quality sequence stop: 587

POLYA=No.

Fax: 301 443 9890

Email: mEST@mail.nih.gov

CDNA Library Preparation: M.B. Soares Lab Clone distribution:
Researchers may obtain BMAP cDNA clones from RESEARCH GENETICS. It
should be noted that Bento Soares is generating a small number of
additional specialized non-redundant arrays of BMAP cDNAs whose
availability will be considered under appropriate and limited
collaborative arrangements

Seq primer: M13 Reverse.

1. 494

/organism="Mus musculus"

/mol_type="mRNA"

/strain="CS7BL/6J"

/db_xref="taxon:10090"

/clone="UI-M-BH0-ake-c-06-0-UI"

/dev_stage="27-32 days"

/lab_host="DH10B (Life Technologies)"

/clone_lib="NIH BMAP_M_S1"

/notes="Vector: pT7T3D-Pac (Pharmacia) with a modified

polylinker; Site 1: Not I; Site 2: Eco RI; The

NIH BMAP_M_S1 library is a subtracted library derived from

a mixture of normalized libraries from ten regions of the

mouse brain (cerebellum, brain stems, olfactory bulbs,

hypothalamus, cortex, amygdala, basal ganglia, pineal

gland, striatum, hippocampus). The driver used for

subtraction consisted of a pool of 20,000 cDNA clones

obtained from non-normalized and normalized libraries of

these ten regions of the mouse brain."

FEATURES
source

Location/Qualifiers
1. .587
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="niaEST:E0488D11-5"
/db_xref="taxon:10090"
/clone="NIA:E0488D11 IMAGE:30877006"
/tissue_type="whole embryo including extraembryonic
tissues at 11.5-days postcoitum"
/dev_stage="E11.5"
/lab_host="DH10B"
/clone_lib="NIA Mouse E11.5 whole embryo cDNA library
(Long)"
/note="Vector: pCMV-SPORT6 (Invitrogen); Site 1: SalI;
Site 2: NotI; Mouse cDNA project by the Laboratory of
Genetics, National Institute on Aging (NIA), Intramural
Research Program, NIH (<http://lgsun.grc.nia.nih.gov/cDNA>).
This is a long-transcript enriched cDNA library (Ref.
Genome Res. 11: 1553-1558 (2001). [PMID: 11544199]). Total
RNAs were extracted from a pool of 3 embryos at 11.5-days
postcoitum. Double-stranded cDNAs were synthesized with an
Oligo(dT) primer [Invitrogen]:
5'-pGACTAGTCTAGATCGGAGCGCCGCTTTT-3' from
2ug of total RNA, treated with T4 DNA polymerase, and
purified by ethanol-precipitation. The cDNAs were ligated
to lone-linker L1-Sal4, purified by phenol/chloroform, and
separated from free linkers by Centricon 100. Then, the
cDNAs were amplified by long-range high fidelity PCR using
Ex Taq polymerase (Takara) with a primer Sal4-S. The
products were purified by phenol/chloroform and Centricion
100. The cDNAs were digested with SalI and NotI enzymes
and cloned into SalI/NotI site of pCMV-SPORT6 plasmid
vector. The DH10B E. coli host was transformed with the
ligation mixture by the standard chemical method. The
average insert size is about 3.3Kb. The library was
constructed by Yulan Piao."

ORIGIN

Query Match 100.0%; Score 25; DB 7; Length 587;
Best Local Similarity 100.0%; Pred. No. 1.9;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAAAACTGGAATCTCAGGCTGAGA 25
|||||
Db 121 AAAAACTGGAATCTCAGGCTGAGA 145
|||||

RESULT 5
CN461714
LOCUS
DEFINITION
CN461714 771 bp mRNA linear EST 21-APR-2004
IMAGE:30650666 5', mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Murinae; Mus.

REFERENCE
1 (bases 1 to 771)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaps-r@mail.nih.gov
Tissue Procurement: Dr. James Lin University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
<http://genome.uiowa.edu/distribution/mousefl.html>

FEATURES
source

This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)
Seq primer: pYX-5
Location/Qualifiers
1. .771
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:30650666"
/tissue_type="Upper Head"
/dev_stage="9.5-10.5 dpc"
/lab_host="DH10B (T1 phage
resistant)"
/clone_lib="NIH BMAP HN0"
/note="Organ: Head; Vector: pYX-Asc; Site 1: EcoR I;
Site 2: Not I; The library was constructed according
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated
with EcoR I adaptor, digested with NotI and then cloned
directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is CCAACTGAAT. This library was created for the University
Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
Developing Mouse Nervous System', supported by National
Institute of Mental Health (NIMH)."

ORIGIN

Query Match 100.0%; Score 25; DB 7; Length 771;
Best Local Similarity 100.0%; Pred. No. 2;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAAAACTGGAATCTCAGGCTGAGA 25
|||||
Db 90 AAAAACTGGAATCTCAGGCTGAGA 114
|||||

RESULT 6
BB268998
LOCUS
DEFINITION
BB268998 808 bp mRNA linear EST 24-OCT-2001
musculus cDNA clone A830037L14 3', mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.

REFERENCE
1 (bases 1 to 808)
Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T.,
Hara, A., Hiramoto, K., Hori, F., Ishii, Y., Ito, M., Kawai, J.,
Konno, H., Kouda, M., Koya, S., Matsuyama, T., Miyazaki, A., Nomura, K.,
Ohno, M., Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F.,
Takeda, Y., Tanaka, T., Toya, T., Muramatsu, M. and Hayashizaki, Y.
RIKEN Mouse ESTs (Arakawa, T., et al. 2001)
Unpublished (2001)

TITLE
JOURNAL
COMMENT
On Jul 7, 2000 this sequence version replaced gi:8965468.
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Sushiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.jp, <http://genome.gsc.riken.jp/>
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.


```

/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:A830037L14"
/db_xref="taxon:10090"
/clone="A830037L14"
/tissue_type="cortex"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="10 days neonate"
1. .1717
misc_feature
/notes="unknown EST (GB|BB641983, evidence: BLASTN, 100%, match=519)"

ORIGIN
Query Match 100.0%; Score 25; DB 4; Length 1717;
Best Local Similarity 100.0%; Pred. No. 2.4;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAACTGGAATCTCAGGCTGAGA 25
|||||
Db 1456 AAAAACTGGAATCTCAGGCTGAGA 1480

RESULT 8
BY618046 395 bp mRNA linear EST 15-DEC-2002
LOCUS BY618046
DEFINITION cDNA clone K330032L03 3', mRNA sequence.
ACCESSION BY618046
VERSION BY618046.1 GI:26953228
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE
1 Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
Nikaido, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I.,
Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,
Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C.,
Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H.,
Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusic, V.,
Chothia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A.,
Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T.,
Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A.,
Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A.,
Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R.,
Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T.,
Numata, K., Okido, T., Pavan, W.J., Perte, G., Pesole, G.,
Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S.,
Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M.,
Sandelin, A., Schneider, C., Semple, C.A., Setou, M., Shimada, K.,
Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M.,
Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y.,
Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I.,
Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P.,
Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M.,
Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,
Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Inotani, K., Ishii, Y.,
Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K.,
Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S.,
Rogers, J., Birney, E. and Hayashizaki, Y.

TITLE Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
PUBMED 12466851
COMMENT Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

```

Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/
 Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S.,
 Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H.,
 Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R.,
 Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K.,
 Shiraki, T., Tagami, M., Waki, K., Watahiki, A., Muramatsu, M. and
 Hayashizaki, Y. Direct Submission
 Computational Analysis of Full-length Mouse cDNAs Compared with
 Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new
 genes. Genome Res. 10 (10), 1617-1630 (2000)
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.
 10 (11), 1757-1771 (2000)
 Computer-based methods for the mouse full-length cDNA
 encyclopedia: real-time sequence clustering for construction of a
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 cDNA library was prepared and sequenced in Mouse Genome
 Encyclopedia Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.
 Division of Experimental Animal Research in Riken contributed to
 prepare mouse tissues.
 Tissues were provided by Michela Fagioli and Takao K. Hensch (
 Laboratory for Neuronal Circuit Development Brain Science Institute
 RIKEN 2-1 Hirotsawa, Wako-shi, Saitama 351-0198 Japan) whose
 assistance we gratefully acknowledge.
 Please visit our web site (http://genome.gsc.riken.go.jp) for
 further details.

FEATURES Location/Qualifiers
 1..395
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="K330032L03"
 /tissue_type="visual cortex"
 /clone_lib="RIKEN full-length enriched, visual cortex"

ORIGIN
 Query Match 96.0%; Score 24; DB 5; Length 395;
 Best Local Similarity 100.0%; Pred. No. 5;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AAAAACTGGAATCTCAGGCTGAGA 25
 |||||
 Db 114 AAAAACTGGAATCTCAGGCTGAGA 137

RESULT 9
 AG325768
 LOCUS AG325768
 DEFINITION Mus musculus molossinus DNA, clone:MSMg01-113L04.T7, genomic survey
 sequence.
 ACCESSION AG325768
 VERSION AG325768.1 GI:47898759
 KEYWORDS GSS.
 SOURCE Mus musculus molossinus (Japanese wild mouse)
 ORGANISM Mus musculus molossinus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
 1 Abe, K., Noguchi, H., Tagawa, K., Yuzuriha, M., Toyoda, A., Kojima, T.,
 Ezawa, K., Saitou, N., Hattori, M., Sakaki, Y., Moriaki, K. and
 Shiroishi, T.
 Contribution of Asian mouse subspecies Mus musculus molossinus to
 genomic constitution of strain C57BL/6J, as defined by BAC-end
 sequence-SNP analysis
 Genome Res. 14 (12), 2439-2447 (2004)
 PUBMED 15574823

/db_xref="taxon:9606"
/clone="IMAGE:4720829"
/lab_host="DH108 (T1 phage-resistant)"
/clone_lib="NIH_MGC_79"
/note="Organ: placenta; Vector: pDNR-LIB (Clontech);
Site 1: SfiI (ggcgctcgcc); Site 2: SfiI
(ggccattatgcc); 5' and 3' adaptors were used in cloning
as follows: 5' adaptor sequence: 5'-CACGCCATTATGCCC-3'
and 3' adaptor sequence:
5'-ATTCTAGAGCGGCGCCGCACATG-dT(30)EN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.3
kb (range 0.5-4.0 kb). 15/15 colonies contained inserts
by PCR. This library was enriched for full-length clones
and was constructed by Clontech Laboratories (Palo Alto,
CA). Note: this is a NIH_MGC Library."

Query Match 80.8%; Score 20.2; DB 2; Length 1069;
Best Local Similarity 88.0%; Pred. No. 3.6e+02;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAACTGGAATCTCAGGCTGAGA 25
|||||
Db 765 AAACAACTGGAATGTCAGGCTGAAA 741

RESULT 16
BF589034/c
LOCUS
DEFINITION 362 bp mRNA linear EST 12-DRC-2000
naa40f10.x1 NCI CGAP Kid11 Homo sapiens cDNA clone IMAGE:3258787
3' mRNA sequence.
BF589034
VERSION BF589034.1 GI:11681358
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 362)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
DNA Sequencing by: Greg Lemon, Ph.D.
CDNA Library Arrayed by: Washington University Genome Sequencing Center
Clone Distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL, send email to:
info@image.llnl.gov
Seq primer: -40UP from Gibco
High quality sequence stop: 294.
Location/Qualifiers
1. .362
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3258787"
/lab_host="DH108"
/clone_lib="NCI CGAP Kid11"
/note="Organ: kidney; Vector: p7T3D-Pac (Pharmacia) with
a modified polylinker; Site 1: Not I; Site 2: Eco RI;
Plasmid DNA from the normalized library NCI CGAP Kid3 was
prepared, and ss circles were made in vitro. Following HAP
purification, this DNA was used as tracer in a subtractive
hybridization reaction. The driver was PCR-amplified cDNAs
from a pool of 5,000 clones made from the same library
(cloneIDs 1322376-1323911, 1456007-1456775, and
1500552-1502855). Subtraction by Bento Soares and M.

Patima Bonaldo. "

Query Match 80.0%; Score 20; DB 2; Length 362;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 AAACCTGGAATCTCAGGCTGA 23
|||||
Db 272 AAACCTGGAATCTCAGGCTGA 253

RESULT 17
BG013717
LOCUS
DEFINITION 460 bp mRNA linear EST 24-JAN-2001
MRI-GN0173-181200-023-b07 GN0173 Homo sapiens cDNA, mRNA sequence.
BG013717
ACCESSION BG013717.1 GI:12464186
VERSION
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 460)
AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V.,
O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
10737800
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=MR1&t2=MR1-GN0173-
181200-023-b07&t3=2000-12-18&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 20
High quality sequence stop: 460.
Location/Qualifiers
1. .460
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="GN0173"
/note="Organ: placenta normal; Vector: puc18; Site 1:
SmaI; Site 2: SmaI; A mini-library was made by cloning
products derived from ORESTES PCR (U.S. Letters Patent
application No. 196,716 - Ludwig Institute for Cancer
Research) profiles into the pUC 18 vector. Reverse
transcription of tissue mRNA and cDNA amplification were
performed under low stringency conditions."

Query Match 80.0%; Score 20; DB 2; Length 460;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 AAACCTGGAATCTCAGGCTGA 23
|||||
Db 204 AAACCTGGAATCTCAGGCTGA 223

RESULT 18
AI479209/c
LOCUS
DEFINITION
tm55904.x1 NCI_CGAP_Kid11 Homo sapiens cDNA clone IMAGE:2162070 3',
mRNA sequence.
ACCESSION
AI479209
VERSION
AI479209.1 GI:4372377
KEYWORDS
EST.
Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
1. (bases 1 to 519)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/nciogap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL
COMMENT
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgaps-x@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 720 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 465.
Location/Qualifiers
1. 519
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2162070"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Kid11"
/note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with
a modified polylinker; Site 1: Not I; Site 2: Eco RI;
Plasmid DNA from the normalized library NCI_CGAP Kid3 was
prepared, and ss circles were made in vitro. Following HAP
purification, this DNA was used as tracer in a subtractive
hybridization reaction. The driver was PCR-amplified cDNAs
from a pool of 5,000 clones made from the same library
(cloneIDs 1322376-1323911, 1456007-1456775, and
1500552-1502855). Subtraction by Bento Soares and M.
Fatima Bonaldo."

ORIGIN

Query Match 80.0%; Score 20; DB 1; Length 519;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 4 AAACCTGGAATCTCAGGCTGA 23
|||||
Db 271 AAACCTGGAATCTCAGGCTGA 252
|||||

RESULT 19
BP250892
LOCUS
DEFINITION
BP250892 Sugano cDNA library, hippocampus Homo sapiens cDNA clone
HPR05485, mRNA sequence.
ACCESSION
BP250892
VERSION
BP250892.1 GI:52133171
KEYWORDS
EST.
Homo sapiens (human)
SOURCE
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
1. (bases 1 to 581)
Suzuki Y., Yanashita R., Shirota M., Sakakibara Y., Chiba J.,
Mizushima-Sugano J., Nakai K. and Sugano S.
Sequence comparison of human and mouse genes reveals a homologous
block structure in the promoter regions
Genome Res. 14 (9), 1711-1718 (2004)
15342556
Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: ysuzuki@ims.u-tokyo.ac.jp.
Location/Qualifiers
1. 581
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="HPR05485"
/tissue_type="hippocampus"
/clone_lib="Sugano cDNA library, hippocampus"

FEATURES

source

ORIGIN

Query Match 80.0%; Score 20; DB 3; Length 581;
Best Local Similarity 100.0%; Pred. No. 3.9e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 4 AAACCTGGAATCTCAGGCTGA 23
|||||
Db 94 AAACCTGGAATCTCAGGCTGA 113
|||||

RESULT 20

EX951384
LOCUS
DEFINITION
DKFZP781E09126_r1 781 (synonym: hlcc4) Homo sapiens cDNA clone
DKFZP781E09126 5', mRNA sequence.
ACCESSION
EX951384
VERSION
EX951384.1 GI:43426263
KEYWORDS
EST.
Homo sapiens (human)
SOURCE
Homo sapiens
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CONTACT

MIPS

INGOLSTAEDTER

LANDSTR.

D-85764

NEUHERBERG

GERMANY

THIS IS THE

5' SEQUENCE

OF THE CLONE

INSERT

CLONE

FROM S. WIEMANN,

MOLECULAR GENOME ANALYSIS,

GERMAN CANCER

RESEARCH CENTER (DKFZ);

EMAIL S.WIEMANN@DKFZ-HEIDELBERG.DE;

SEQUENCED BY DKFZ (GERMAN CANCER RESEARCH CENTER,

HEIDELBERG/GERMANY) WITHIN THE CDNA SEQUENCING CONSORTIUM OF THE

GERMAN GENOME PROJECT.

NO SI SEQUENCE AVAILABLE.

THIS CLONE (DKFZP781E09126) IS AVAILABLE AT THE RZPD IN BERLIN.

PLEASE CONTACT THE RZPD: RESSOURCENZENTRUM, HEUBERWEG 6, 14059

BERLIN-CHARLOTTENBURG, GERMANY; EMAIL: clone@rzpd.de.

LOCATION/QUALIFIERS

1. 767

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="DKFZP781E09126"

/dev_stage="adult"

```

/lab_host="DH10B"
/clone_lib="781 (synonym: hlcc4)"
/note="Vector: pSport1_Sfi; Site_1: SfiIA; Site_2: SfiIB;
cDNA-collection"

ORIGIN
Query Match      80.0%; Score 20; DB 5; Length 767;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 AAACCTGGAATCTCAGGCTGA 23
    |||||
Db 708 AAACCTGGAATCTCAGGCTGA 727

RESULT 21
AQ378916
LOCUS
DEFINITION
  AQ378916 661 bp DNA linear GSS 20-MAY-1999
  genomic survey sequence.
ACCESSION
  AQ378916
VERSION
  AQ378916.1 GI:4349939
KEYWORDS
  GSS.
SOURCE
  Homo sapiens (human)
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
  Homnidae; Homo.
REFERENCE
  1 (bases 1 to 661)
  Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and
  Venter,J.C
  Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
  Map Building
  Unpublished (1997)
JOURNAL
  Other GSSs: RPCI11-164114.TV
COMMENT
  Contact: Shaying Zhao, William Nierman, Mark Adams
  Department of Eukaryotic Genomics
  The Institute for Genomic Research
  9712 Medical Center Dr., Rockville, MD 20850
  Tel: 301 838 0200
  Fax: 301 838 0208
  Email: hbe@tigr.org
  Clones are derived from the human BAC library RPCI-11. For BAC
  library availability, please contact Pieter de Jong
  (pietere@jmg.med.buffalo.edu). Clones may be purchased from
  BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
  Research Genetics (info@resgen.com). BAC end search page:
  http://www.tigr.org/cdb/humgen/bac_end_search/bac_end_search.html
  Seq primer: SP6
  Class: BAC ends.
FEATURES
  source
    Location/Qualifiers
      1..661
        /organism="Homo sapiens"
        /mol_type="genomic DNA"
        /db_xref="GDB:7562797"
        /db_xref="taxon:9606"
        /clone="RPCI-11-164114"
        /sex="Male"
        /cell_type="Lymphocytes"
        /clone_lib="RPCI-11"
        /note="Vector: pBACe3.6; Site 1: EcoRI; Site_2: EcoRI;
        RPCI11 Human Male BAC Library"

ORIGIN
Query Match      79.2%; Score 19.8; DB 9; Length 661;
Best Local Similarity 91.3%; Pred. No. 4.9e+02;
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AAAAACTGGAATCTCAGGCTGAG 24
    |||||
Db 550 AGAATCTGGAATCTCAGGCTGAG 572

```

```

RESULT 22
BP738553/c
LOCUS
DEFINITION
  BP738553 Osada Taira anterior neuroectoderm (ANE) PCS105 cDNA
  library Xenopus laevis cDNA clone XL516n04ex 3', mRNA sequence.
ACCESSION
  BP738553
VERSION
  BP738553.1 GI:46087146
KEYWORDS
  EST.
SOURCE
  Xenopus laevis (African clawed frog)
  Xenopus laevis
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
  Xenopodinae; Xenopus; Xenopus.
REFERENCE
  1 (bases 1 to 685)
  Osada,S., Kitayama,A., Ueno,N. and Taira,M.
  Expression analysis of genes which are expressed in the anterior
  neuroectoderm of Xenopus embryos
  Unpublished (2004)
JOURNAL
  Contact: Masanori Taira
  Department of Biological Sciences
  Graduate School of Science, University of Tokyo; CREST, Japan
  Science and Technology Corporation, Japan
  7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan
  Tel: 81-03-5841-4434
  Fax: 81-03-5841-4434
  Email: m.taira@biol.s.u-tokyo.ac.jp,
  URL:http://www.shigen.nig.ac.jp/abrp/xenopus/est/.
  Location/Qualifiers
    1..685
      /organism="Xenopus laevis"
      /mol_type="mRNA"
      /db_xref="taxon:8355"
      /clone="XL516n04ex"
      /tissue_type="anterior neuroectoderm"
      /dev_stage="late gastrula (stage 12.5)"
      /clone_lib="Osada Taira anterior neuroectoderm (ANE)
      PCS105 cDNA library"

ORIGIN
Query Match      79.2%; Score 19.8; DB 3; Length 685;
Best Local Similarity 91.3%; Pred. No. 5e+02;
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAACTGGAATCTCAGGCTGA 23
    |||||
Db 492 AAAAAACGGAATCTCAGGATGA 470

RESULT 23
CO932221
LOCUS
DEFINITION
  CO932221 898 bp mRNA linear EST 16-AUG-2004
  AGENCOURT_30436848 NIH_ZGC_14 Danio rerio cDNA clone IMAGE:7406857
  5', mRNA sequence.
ACCESSION
  CO932221
VERSION
  CO932221.1 GI:51286898
KEYWORDS
  EST.
SOURCE
  Danio rerio (zebrafish)
  Danio rerio
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
  Cypriniformes; Cyprinidae; Danio.
REFERENCE
  1 (bases 1 to 898)
  NIH-MGC http://mgc.nci.nih.gov/.
  National Institutes of Health, Mammalian Gene Collection (MGC)
  Unpublished (1999)
  Contact: Daniela S. Gerhard, Ph.D.
  Office of Cancer Genomics / NIH
  National Cancer Institute
  Bldg. 31 Rm10A07 Bethesda, MD 20892
  Email: cgapbs-r@mail.nih.gov
  Tissue Procurement: John Ngai, Nancy Freeman, NIDCD
  cDNA Library Preparation: Dr. Sumio Sugano
  cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

```

DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LHAM15607 row: d column: 23
High quality sequence start: 10
High quality sequence stop: 351.
Location/Qualifiers

FEATURES
source

```
1. .898
/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/clone="IMAGE:7406857"
/tissue_type="olfactory epithelium"
/lab_host="DH10B Tona"
/clone_lib="NIH_ZGC_14"
/notes="Organ: olfactory epithelium; Vector: pME18S-FL3;
Site: 1: DraIII; Site 2: DraIII; 1st strand cDNA was primed
with an oligo(dT) primer
[GGGCTGAAGCGGCTATGTGGCTTTTTTTTTTTTTTTT];
double-stranded cDNA was ligated to a DraIII adaptor
[GGCTTACUGG], digested and directionally cloned into
distinct DraIII sites of the pME18S-FL3. Library was size
selected for 1.0 kb, with a average insert size of ~1.2kb.
Library constructed by Yutaka Suzuki (University of Tokyo
Institute of Medical Science). Custom primers recommended
for sequencing: 5' end primer 5'-GGATGTTCCTTACTCTA-3'
and 3' end primer 5'-CGACTGAGTCGAGACA-3'. Note: This
is a Zebrafish Gene Collection (ZGC) library"
```

ORIGIN

Query Match 77.6%; Score 19.4; DB 7; Length 898;
Best Local Similarity 95.2%; Pred. No. 8e+02;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AAAAACTGGAATCTCAGGCTG 22
||| ||||| ||||| ||||| |||||
Db 698 AAGAACTGGAATCTCAGGCTG 718

RESULT 24

AV169292/c
LOCUS AV169292 Mus musculus head C57BL/6J 13-day embryo Mus musculus cDNA
DEFINITION Clone 3110083I21, mRNA sequence.

ACCESSION AV169292
VERSION AV169292.1 GI:5375729
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 (bases 1 to 317)
Carninci, P., Shibata, K., Ozawa, Y., Konno, H., Itoh, M., Aizawa, K.,
Akahira, S., Akiyama, J., Fukuda, S., Fukunishi, Y., Funayama, T.,
Hara, A., Hayatsu, N., Hori, F., Ishikawa, T., Itoh, M., Izawa, M.,
Kawai, J., Kikuchi, N., Kojima, Y., Matsuyama, T., Niitsuma, H., Oda, H.,
Owa, C., Sato, K., Shibata, Y., Shigemoto, Y., Shiraki, T., Sogabe, Y.,
Sugahara, Y., Suzuki, H., Suzuki, H., Tateno, M., Tomaru, Y.,
Tomimaga, N., Watanabe, S., Yagame, M., Yamamura, T., Yokota, T.,
Yoshino, M., Muramatsu, M., Okazaki, Y. and Hayashizaki, Y.
RIKEN Mouse ESTs
Unpublished (1999)
Contact: Chie Owa
Genome Science Laboratory

TITLE
JOURNAL
COMMENT

3-1-1 Koyada, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-9145
Fax: 81-298-36-9098
Email: genome-res@rtc.riken.go.jp
Thermotabilization and thermoactivation of thermolabile enzymes by
trehalose and its application for the synthesis of full length cDNA

(Proc. Natl. Acad. Sci. U.S.A. 95(2):520-524 (1998))
Transcriptional sequencing: A method for DNA sequencing using RNA
polymerase (Proc. Natl. Acad. Sci. U.S.A. 95(7):3455-3460 (1998))
Please visit our web site (<http://genome.rtc.riken.go.jp>) for
further details.

FEATURES
source

```
Location/Qualifiers
1. .317
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="3110083I21"
/sex="mixed"
/tissue_type="head"
/dev_stage="13-day embryo"
/clone_lib="Mus musculus head C57BL/6J 13-day embryo"
```

ORIGIN

Query Match 76.8%; Score 19.2; DB 1; Length 317;
Best Local Similarity 87.5%; Pred. No. 8e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAACTGGAATCTCAGGCTGAG 24
||| ||||| ||||| ||||| |||||
Db 117 AAAAACTGGAATCTCAGGCGGAG 94

RESULT 25

BP070142
LOCUS BP070142 Lotus corniculatus var. japonicus nodulating root Lotus
DEFINITION corniculatus var. japonicus cDNA clone GNF038d05 3', mRNA sequence.

ACCESSION BP070142
VERSION BP070142.1 GI:45626868
KEYWORDS EST.
SOURCE Lotus corniculatus var. japonicus (Lotus japonicus)
ORGANISM Lotus corniculatus var. japonicus

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Lotaeae;
Lotus.

REFERENCE

1 (bases 1 to 410)
Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S.
Characteristics of the Lotus japonicus gene repertoire deduced from
large-scale expressed sequence tag (EST) analysis
Plant Mol. Biol. 54 (3), 405-414 (2004)

COMMENT

Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL: <http://www.kazusa.or.jp/en/plant/>.

FEATURES
source

```
Location/Qualifiers
1. .410
/organism="Lotus corniculatus var. japonicus"
/mol_type="mRNA"
/isolate="Gifu B-129"
/db_xref="taxon:34305"
/clone="GNF038d05"
/tissue_type="nodulating root"
/clone_lib="Lotus corniculatus var. japonicus nodulating
root"
/notes="root of hypernodulating mutant harl-3, rich in
nodule primordia"
```

ORIGIN

Query Match 76.8%; Score 19.2; DB 3; Length 410;
Best Local Similarity 87.5%; Pred. No. 8.5e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAACTGGAATCTCAGGCTGAGA 25
||| ||||| ||||| ||||| |||||
Db 156 AAAAACTGGAATCTGAGCAAGA 179

```

RESULT 26
CA078973
LOCUS          535 bp      mRNA      linear      EST 23-SEP-2003
DEFINITION    SCRLAM1013F07.g AM1 Saccharum officinarum cDNA clone SCRLAM1013F07
5', mRNA sequence.
ACCESSION    CA078973
VERSION      CA078973.1  GI:34931245
KEYWORDS     EST.
SOURCE       Saccharum officinarum
ORGANISM     Saccharum officinarum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
Clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum
complex.
REFERENCE    1 (bases 1 to 535)
AUTHORS      Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
TITLE        The libraries that made SUCEST
JOURNAL      Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
COMMENT      Contact: Arruda P
Centro de Biologia Molecular e Engenhariaia Genetica
Universidade Estadual de Campinas
Caixa Postal 6010, 13083-970, Campinas SP, Brazil
Tel: 55 19 3788 1137
Fax: 55 19 3788 1089
Email: parruda@unicamp.br
Clone distribution: Clone distribution information can be found
through the Brazilian Clone Collection Center (BCCC) at
http://www.bcccenter.fcav.unesp.br
Plate: 013 row: F column: 07
Seq primer: T7 Promoter Primer.
FEATURES     source
             Location/Qualifiers
             1..535
                /organism="Saccharum officinarum"
                /mol_type="mRNA"
                /db_xref="taxon:4547"
                /clone="SCRLAM1013F07"
                /lab_host="DH10B"
                /clone_lib="AM1"
                /notes="Organ: Apical meristem and tissues surrounding of
                mature plants; Vector: pSport1; Site 1: SalI; Site 2:
                NotI; An unidirectional cDNA library generated from
                [Apical meristem and tissues surrounding of mature
                plant]. cDNA was prepared from polyA+ mRNA using
                SuperScript Plasmid System Kit (Invitrogen). The
                double-strand cDNAs were fractionated in a sepharose
                CL-2B 40cm-columns and fragments sizing between 0.8 and
                1.5 Kb were directionally cloned into the vector. Details
                of each source of RNA and library construction can be
                obtained at http://sucest.lad.ic.unicamp.br/public"

ORIGIN
Query Match          76.8%; Score 19.2; DB 6; Length 535;
Best Local Similarity 87.5%; Pred. No. 8.9e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY  2  AAAAATCGGAATCTCAGGCTGAGA 25
      |||||
Db   376 AAAAGCTGGAATCTCAAGCTGAAA 399

RESULT 27
BP058178
LOCUS          562 bp      mRNA      linear      EST 25-AUG-2004
DEFINITION    BP058178 Lotus corniculatus var. japonicus cDNA clone SPDL098f12_f
2', mRNA sequence.
ACCESSION    BP058178
VERSION      BP058178.1  GI:45614929
KEYWORDS     EST.
SOURCE       Lotus corniculatus var. japonicus
ORGANISM     Lotus corniculatus var. japonicus

```

```

REFERENCE    1 (bases 1 to 562)
AUTHORS      Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.
TITLE        Characteristics of the Lotus japonicus gene repertoire deduced from
large-scale expressed sequence tag (EST) analysis
JOURNAL      Plant Mol. Biol. 54 (3), 405-414 (2004)
PUBMED       15284495
COMMENT      Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.
Location/Qualifiers
             1..562
                /organism="Lotus corniculatus var. japonicus"
                /mol_type="mRNA"
                /isolate="Miyakojima MG-20"
                /db_xref="taxon:34305"
                /clone="SPDL098f12_f"
                /tissue_type="pods"
                /dev_stages="pod (less than 20 mm in length)"
                /clone_lib="Lotus corniculatus var. japonicus pods (less
                than 20 mm in length)"

ORIGIN
Query Match          76.8%; Score 19.2; DB 3; Length 562;
Best Local Similarity 87.5%; Pred. No. 9e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY  2  AAAAATCGGAATCTCAGGCTGAGA 25
      |||||
Db   153 AAAAATCGGAATCTGAGGCAAGA 176

RESULT 28
CR868598/c
LOCUS          637 bp      DNA      linear      GSS 19-NOV-2004
DEFINITION    Sus scrofa BES, genomic survey sequence.
ACCESSION    CR868598
VERSION      CR868598.1  GI:55866844
KEYWORDS     GSS; Bac-end sequence BES; Genome Survey Sequence.
SOURCE       Sus scrofa (pig)
ORGANISM     Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae;
Sus.
REFERENCE    1 (bases 1 to 637)
AUTHORS      Rogel-Gaillard,C., Bourgeaux,N., Billault,A., Vaiman,M. and
Chardon,P.
TITLE        Construction of a swine BAC library: application to the
characterization and mapping of porcine type C endoviral elements
JOURNAL      Cytogenet. Cell Genet. 85 (3-4), 205-211 (1999)
PUBMED       10449899
REFERENCE    2 (bases 1 to 637)
AUTHORS      Chardon,P., Iannuccelli,N., Roig,A., Dossat,C., Demars,J.,
Rogel-Gaillard,C., Roy,A., Schibler,L. and Milan,D.
TITLE        A physical map of the swine genome
JOURNAL      Unpublished
REFERENCE    3 (bases 1 to 637)
AUTHORS      Genoscope.
TITLE        Direct Submission
JOURNAL      Submitted (18-NOV-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr)
- Web : www.genoscope.cns.fr
Location/Qualifiers
             1..637
                /organism="Sus scrofa"
                /mol_type="genomic DNA"
                /strain="Large White"
                /db_xref="taxon:9823"

```

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Lotaeae; Lotus.

[illegible]

the JGI Clone Id and the direction of sequencing. The suffix '.rev' indicates a reverse sequencing read of the insert. It does not necessarily reflect the orientation of the insert.
 Poly-A: Based upon the presence of a run of 14 or more T residues at the beginning of the sequence, this clone was polyadenylated.
 The resulting Poly-T sequence has been removed.
 Small Insert: Based upon one or more sequencing reads of this clone where vector sequence was present at both ends, this clone has been determined to contain a cDNA insert on the order of 600-1000 bases.
 Plate: ANNN 0089 row: a column: 7
 High quality sequence stop: 556
 POLYA=Yes.

FEATURES

Location/Qualifiers
 1..682
 /organism="Pimephales promelas"
 /mol_type="mRNA"
 /db_xref="taxon:90988"
 /clone="ANNN8473"
 /tissue_type="Whole"
 /dev_stage="Adult"
 /clone_lib="ANNN Pimephales promelas Whole (L)"
 /note="Vector: pDONR222; The library was made from dT primed cDNA and cloned into Invitrogen vector pDONR222. Poly A RNA were primed with a Biotin-attB2-Oligo(dT) primer (5'- GGCGCGCGCACAACTTTGTACAGAAAGTTGGGT(T)19 -3') and the first strand synthesized using Superscript II (Invitrogen). The second strand was synthesized and an attB1 adaptor was ligated to the 5' end (5'- TCCTCGGGGCAACATTGTACAAAAGTTGCC -3') . cDNA was size selected using 1% agarose gel electrophoresis (L ~0.5-1.2k, M ~1.2-2.5k, H ~2.5k) and then inserted into the vector using site specific recombination (flanking attB sites on cDNA). The work was done at DOE Joint Genome Institute."

source
 1..693
 /organism="Pimephales promelas"
 /mol_type="mRNA"
 /db_xref="taxon:90988"
 /clone="ANNN7222"
 /tissue_type="Whole"
 /dev_stage="Adult"
 /clone_lib="ANNN Pimephales promelas Whole (L)"
 /note="Vector: pDONR222; The library was made from dT primed cDNA and cloned into Invitrogen vector pDONR222. Poly A RNA were primed with a Biotin-attB2-Oligo(dT) primer (5'- GGCGCGCGCACAACTTTGTACAGAAAGTTGGGT(T)19 -3') and the first strand synthesized using Superscript II (Invitrogen). The second strand was synthesized and an attB1 adaptor was ligated to the 5' end (5'- TCCTCGGGGCAACATTGTACAAAAGTTGCC -3') . cDNA was size selected using 1% agarose gel electrophoresis (L ~0.5-1.2k, M ~1.2-2.5k, H ~2.5k) and then inserted into the vector using site specific recombination (flanking attB sites on cDNA). The work was done at DOE Joint Genome Institute."

ORIGIN

Query Match 76.8%; Score 19.2; DB 8; Length 682;
 Best Local Similarity 87.5%; Pred. No. 9.4e+02;
 Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAACTGGAATCTCAGGCTGAGA 25
 |||||
 Db 604 AAAAACTGCAATATCAGCCTGAGA 627

RESULT 33
 DT105042
 LOCUS
 DEFINITION
 JGI ANNN7222.rev ANNN Pimephales promelas Whole (L) Pimephales promelas cDNA clone ANNN7222 3', mRNA sequence.
 DT105042
 DT105042.1 GI:72455702
 EST.

ORGANISM
 Pimephales promelas
 Pimephales promelas
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Pimephales.
 1 (bases 1 to 693)
 Richardson,P., Lucas,S., Rokhsar,D., Detter,J.C., Ng,D.C., Brokstein,P. and Lindquist,E.A.
 DOE Joint Genome Institute Pimephales promelas EST project
 Other ESTs: JGI ANNN7222.fwd
 Contact: Lindquist,E.A., Richardson,P.
 DOE Joint Genome Institute
 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
 Tel: 925 296 5600
 Fax: 925 296 5710
 Email: cdna@jgi-psf.org
 cDNA Library Preparation: DOE Joint Genome Institute:
 http://www.jgi.doe.gov

REFERENCE
 AUTHORS
 Richardson,P., Lucas,S., Rokhsar,D., Detter,J.C., Ng,D.C., Brokstein,P. and Lindquist,E.A.
 DOE Joint Genome Institute Pimephales promelas EST project
 Other ESTs: JGI ANNN7222.fwd
 Contact: Lindquist,E.A., Richardson,P.
 DOE Joint Genome Institute
 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
 Tel: 925 296 5600
 Fax: 925 296 5710
 Email: cdna@jgi-psf.org
 cDNA Library Preparation: DOE Joint Genome Institute:
 http://www.jgi.doe.gov

TITLE
 JOURNAL
 COMMENT

ORIGIN

Query Match 76.8%; Score 19.2; DB 8; Length 693;
 Best Local Similarity 87.5%; Pred. No. 9.4e+02;
 Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAACTGGAATCTCAGGCTGAGA 25
 |||||
 Db 604 AAAAACTGCAATATCAGCCTGAGA 627

RESULT 34
 DT110690/c
 LOCUS
 DEFINITION

JGI ANNO998.fwd ANNO Pimephales promelas Whole (M) Pimephales promelas cDNA clone ANNO998 5', mRNA sequence.
 DT110690
 DT110690.1 GI:72467281
 EST.

ORGANISM
 Pimephales promelas
 Pimephales promelas
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Pimephales.
 1 (bases 1 to 694)
 Richardson,P., Lucas,S., Rokhsar,D., Detter,J.C., Ng,D.C., Brokstein,P. and Lindquist,E.A.
 DOE Joint Genome Institute Pimephales promelas EST project
 Other ESTs: JGI ANNO998.rev
 Contact: Lindquist,E.A., Richardson,P.
 DOE Joint Genome Institute
 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
 Tel: 925 296 5600
 Fax: 925 296 5710
 Email: cdna@jgi-psf.org
 cDNA Library Preparation: DOE Joint Genome Institute:
 http://www.jgi.doe.gov

REFERENCE
 AUTHORS
 Richardson,P., Lucas,S., Rokhsar,D., Detter,J.C., Ng,D.C., Brokstein,P. and Lindquist,E.A.
 DOE Joint Genome Institute Pimephales promelas EST project
 Other ESTs: JGI ANNO998.rev
 Contact: Lindquist,E.A., Richardson,P.
 DOE Joint Genome Institute
 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
 Tel: 925 296 5600
 Fax: 925 296 5710
 Email: cdna@jgi-psf.org
 cDNA Library Preparation: DOE Joint Genome Institute:
 http://www.jgi.doe.gov

TITLE
 JOURNAL
 COMMENT

Naming Conventions: EST name is generated by the concatenation of the JGI Clone Id and the direction of sequencing. The suffix '.rev' indicates a reverse sequencing read of the insert. It does not necessarily reflect the orientation of the insert.
 Poly-A: Based upon the presence of a run of 14 or more T residues at the beginning of the sequence, this clone was polyadenylated.
 The resulting Poly-T sequence has been removed.
 Small Insert: Based upon one or more sequencing reads of this clone where vector sequence was present at both ends, this clone has been determined to contain a cDNA insert on the order of 600-1000 bases.
 Plate: ANNN 0073 row: 1 column: 6
 High quality sequence stop: 554
 POLYA=Yes.

Location/Qualifiers
 1..693
 /organism="Pimephales promelas"
 /mol_type="mRNA"
 /db_xref="taxon:90988"
 /clone="ANNN7222"
 /tissue_type="Whole"
 /dev_stage="Adult"
 /clone_lib="ANNN Pimephales promelas Whole (L)"
 /note="Vector: pDONR222; The library was made from dT primed cDNA and cloned into Invitrogen vector pDONR222. Poly A RNA were primed with a Biotin-attB2-Oligo(dT) primer (5'- GGCGCGCGCACAACTTTGTACAGAAAGTTGGGT(T)19 -3') and the first strand synthesized using Superscript II (Invitrogen). The second strand was synthesized and an attB1 adaptor was ligated to the 5' end (5'- TCCTCGGGGCAACATTGTACAAAAGTTGCC -3') . cDNA was size selected using 1% agarose gel electrophoresis (L ~0.5-1.2k, M ~1.2-2.5k, H ~2.5k) and then inserted into the vector using site specific recombination (flanking attB sites on cDNA). The work was done at DOE Joint Genome Institute."

source
 1..693
 /organism="Pimephales promelas"
 /mol_type="mRNA"
 /db_xref="taxon:90988"
 /clone="ANNN7222"
 /tissue_type="Whole"
 /dev_stage="Adult"
 /clone_lib="ANNN Pimephales promelas Whole (L)"
 /note="Vector: pDONR222; The library was made from dT primed cDNA and cloned into Invitrogen vector pDONR222. Poly A RNA were primed with a Biotin-attB2-Oligo(dT) primer (5'- GGCGCGCGCACAACTTTGTACAGAAAGTTGGGT(T)19 -3') and the first strand synthesized using Superscript II (Invitrogen). The second strand was synthesized and an attB1 adaptor was ligated to the 5' end (5'- TCCTCGGGGCAACATTGTACAAAAGTTGCC -3') . cDNA was size selected using 1% agarose gel electrophoresis (L ~0.5-1.2k, M ~1.2-2.5k, H ~2.5k) and then inserted into the vector using site specific recombination (flanking attB sites on cDNA). The work was done at DOE Joint Genome Institute."

DNA Sequencing: DOE Joint Genome Institute: <http://www.jgi.doe.gov>
 Naming Conventions: EST name is generated by the concatenation of the JGI Clone Id and the direction of sequencing. The suffix '.fwd' indicates a forward sequencing read of the insert. It does not necessarily reflect the orientation of the insert.
 Small Insert: Based upon one or more sequencing reads of this clone where vector sequence was present at both ends, this clone has been determined to contain a cDNA insert on the order of 600-1000 bases.
 Plate: ANNO 0009 row: k column: 10
 High quality sequence stop: 681.

FEATURES

source

Location/Qualifiers

1..694

/organism="Pimephales promelas"

/mol_type="mRNA"

/db_xref="taxon:90988"

/clone="ANNO998"

/tissue_type="Whole"

/dev_stage="Adult"

/clone_lib="ANNO Pimephales promelas Whole (M)"

/note="Vector: pDONR222; The library was made from dT

primed cDNA and cloned into Invitrogen vector pDONR222.

Poly A RNA were primed with a Biotin-attB2-Oligo(dT)

primer (5'-GGCGGCGCACAACTTTGTACAGAAAGTTGGT(19)-3')

and the first strand synthesized using Superscript II

(Invitrogen). The second strand was synthesized and an

attB1 adaptor was ligated to the 5' end (5'-

TCGTCGGGACAACTTTGTACAGAAAGTTGG-3'). The cDNA was size

selected using 1% agarose gel electrophoresis. (L

-0.5-1.2k, M -1.2-2.5k, H -2.5k) and then inserted into

the vector using site specific recombination (flanking

attB sites on cDNA). The work was done at DOE Joint Genome

Institute."

ORIGIN

Query Match 76.8%; Score 19.2; DB 8; Length 694;

Best Local Similarity 87.5%; Pred. No. 9.4e+02;

Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAACTGGAATCTCAGGCTGAGA 25

||||| ||||| ||||| ||||| |||||

Db 144 AAAAACTGCAATATCAGCTGAGA 121

RESULT 35

Dt118130/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Pimephales promelas

Pimephales promelas

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Actinopterygii; Neopterygii; Teleostei; Ostariophysi;

Cypriniformes; Cyprinidae; Pimephales.

1 (bases 1 to 705)

Richardson, P., Lucas, S., Rokhsar, D., Dettler, J.C., Ng, D.C.,

Brokstein, P. and Lindquist, E.A.

DOE Joint Genome Institute Pimephales promelas EST project

Unpublished (2005)

Other ESTs: JGI ANNO5327.rev

Contact: Lindquist, E.A., Richardson, P.

DOE Joint Genome Institute

2800 Mitchell Drive, Walnut Creek, CA 94598, USA

Tel: 925 296 5600

Fax: 925 296 5710

Email: cdna@jgi-psf.org

cDNA Library Preparation: DOE Joint Genome Institute:

<http://www.jgi.doe.gov>

DNA Sequencing: DOE Joint Genome Institute: <http://www.jgi.doe.gov>

Naming Conventions: EST name is generated by the concatenation of

the JGI Clone Id and the direction of sequencing. The suffix '.fwd' indicates a forward sequencing read of the insert. It does not necessarily reflect the orientation of the insert.
 Small Insert: Based upon one or more sequencing reads of this clone where vector sequence was present at both ends, this clone has been determined to contain a cDNA insert on the order of 600-1000 bases.
 Plate: ANNO 0053 row: n column: 12
 High quality sequence stop: 624.

FEATURES

source

Location/Qualifiers

1..705

/organism="Pimephales promelas"

/mol_type="mRNA"

/db_xref="taxon:90988"

/clone="ANNO5327"

/tissue_type="Whole"

/dev_stage="Adult"

/clone_lib="ANNO Pimephales promelas Whole (M)"

/note="Vector: pDONR222; The library was made from dT

primed cDNA and cloned into Invitrogen vector pDONR222.

Poly A RNA were primed with a Biotin-attB2-Oligo(dT)

primer (5'-GGCGGCGCACAACTTTGTACAGAAAGTTGGT(19)-3')

and the first strand synthesized using Superscript II

(Invitrogen). The second strand was synthesized and an

attB1 adaptor was ligated to the 5' end (5'-

TCGTCGGGACAACTTTGTACAGAAAGTTGG-3'). The cDNA was size

selected using 1% agarose gel electrophoresis. (L

-0.5-1.2k, M -1.2-2.5k, H -2.5k) and then inserted into

the vector using site specific recombination (flanking

attB sites on cDNA). The work was done at DOE Joint Genome

Institute."

ORIGIN

Query Match 76.8%; Score 19.2; DB 8; Length 705;

Best Local Similarity 87.5%; Pred. No. 9.5e+02;

Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAACTGGAATCTCAGGCTGAGA 25

||||| ||||| ||||| ||||| |||||

Db 128 AAAAACTGCAATATCAGCCTGAGA 105

RESULT 36

Dt100080/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Pimephales promelas

Pimephales promelas

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Actinopterygii; Neopterygii; Teleostei; Ostariophysi;

Cypriniformes; Cyprinidae; Pimephales.

1 (bases 1 to 706)

Richardson, P., Lucas, S., Rokhsar, D., Dettler, J.C., Ng, D.C.,

Brokstein, P. and Lindquist, E.A.

DOE Joint Genome Institute Pimephales promelas EST project

Unpublished (2005)

Other ESTs: JGI ANNA4389.rev

Contact: Lindquist, E.A., Richardson, P.

DOE Joint Genome Institute

2800 Mitchell Drive, Walnut Creek, CA 94598, USA

Tel: 925 296 5600

Fax: 925 296 5710

Email: cdna@jgi-psf.org

cDNA Library Preparation: DOE Joint Genome Institute:

<http://www.jgi.doe.gov>

DNA Sequencing: DOE Joint Genome Institute: <http://www.jgi.doe.gov>

Naming Conventions: EST name is generated by the concatenation of

FEATURES
source

Location/Qualifiers

1..738
/organism="Pimephales promelas"
/mol_type="mRNA"
/db_xref="taxon:90988"
/clone="ANN7528"
/tissue_type="Whole"
/dev_stage="Adult"
/clone_lib="ANN Pimephales promelas Whole (L)"
/note="Vector: pDONR222; The library was made from dT primed cDNA and cloned into Invitrogen vector pDONR222. Poly A RNA were primed with a Biotin-attB2-Oligo(dT) primer (5'-GGCGCCGACAACTTTGTACAGAAAGTTGGT(T)19-3') and the first strand synthesized using Superscript II (Invitrogen). The second strand was synthesized and an attB1 adaptor was ligated to the 5' end (5'-TCGTCGGGACAACTTTGTACAAAAAGTTGG-3' and 5'-CCACTTTTGTACAAAGTTGTCCCC-3'). cDNA was size selected using 1% agarose gel electrophoresis (L -0.5-1.2K, M -1.2-2.5k, H ->2.5k) and then inserted into the vector using site specific recombination (flanking attB sites on cDNA). The work was done at DOE Joint Genome Institute."

ORIGIN

Query Match 76.8%; Score 19.2; DB 8; Length 738;
Best Local Similarity 87.5%; Pred. No. 9.6e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAACTGGAATCTCAGGCTGAGA 25

|||||
Db 166 AAAAACTGCAATATCAGCCTGAGA 143

RESULT 39

DT108879/c

LOCUS

DEFINITION

Promelas cDNA clone ANNN11096 5', mRNA sequence.

ACCESSION

DT108879

VERSION

DT108879.1

KEYWORDS

Pimephales promelas

SOURCE

Pimephales promelas

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Actinopterygii; Neopterygii; Teleostei; Ostariophysi;

Cypriniformes; Cyprinidae; Pimephales.

1 (bases 1 to 743)

Richardson, P., Lucas, S., Rokhsar, D., Dettter, J.C., Ng, D.C.,

Brokstein, P., and Lindquist, E.A.

DOE Joint Genome Institute Pimephales promelas EST project

Unpublished (2005)

Contact: Lindquist, E.A., Richardson, P.

DOE Joint Genome Institute

2800 Mitchell Drive, Walnut Creek, CA 94598, USA

Tel: 925 296 5600

Fax: 925 296 5710

Email: cdna@jgi-psf.org

cDNA Library Preparation: DOE Joint Genome Institute:

http://www.jgi.doe.gov

DNA Sequencing: DOE Joint Genome Institute: http://www.jgi.doe.gov

Naming Conventions: EST name is generated by the concatenation of

the JGI Clone Id and the direction of sequencing. The suffix '.fwd'

indicates a forward sequencing read of the insert. It does not

necessarily reflect the orientation of the insert.

Small Insert: Based upon one or more sequencing reads of this clone

where vector sequence was present at both ends, this clone has been

determined to contain a cDNA insert on the order of 600-1000 bases.

Plate: ANNN 0113 row: p column: 14

High quality sequence stop: 730.

Location/Qualifiers

1..743

/organism="Pimephales promelas"

/mol_type="mRNA"

/db_xref="taxon:90988"

/clone="ANN11096"

/tissue_type="Whole"

/dev_stage="Adult"

/clone_lib="ANN Pimephales promelas Whole (L)"

/note="Vector: pDONR222; The library was made from dT

primed cDNA and cloned into Invitrogen vector pDONR222.

Poly A RNA were primed with a Biotin-attB2-Oligo(dT)

primer (5'-GGCGCCGACAACTTTGTACAGAAAGTTGGT(T)19-3')

and the first strand synthesized using Superscript II

(Invitrogen). The second strand was synthesized and an

attB1 adaptor was ligated to the 5' end (5'-

TCGTCGGGACAACTTTGTACAAAAAGTTGG-3' and 5'-

CCACTTTTGTACAAAGTTGTCCCC-3'). cDNA was size selected

using 1% agarose gel electrophoresis (L -0.5-1.2K, M

-1.2-2.5k, H ->2.5k) and then inserted into the vector

using site specific recombination (flanking attB sites on

cDNA). The work was done at DOE Joint Genome Institute."

ORIGIN

Query Match 76.8%; Score 19.2; DB 8; Length 743;
Best Local Similarity 87.5%; Pred. No. 9.6e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAACTGGAATCTCAGGCTGAGA 25

|||||
Db 141 AAAAACTGCAATATCAGCCTGAGA 118

RESULT 40

DT093604/c

LOCUS

DEFINITION

Promelas cDNA clone ANNN725 5', mRNA sequence.

ACCESSION

DT093604

VERSION

DT093604.1

KEYWORDS

Pimephales promelas

SOURCE

Pimephales promelas

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Actinopterygii; Neopterygii; Teleostei; Ostariophysi;

Cypriniformes; Cyprinidae; Pimephales.

1 (bases 1 to 755)

Richardson, P., Lucas, S., Rokhsar, D., Dettter, J.C., Ng, D.C.,

Brokstein, P., and Lindquist, E.A.

DOE Joint Genome Institute Pimephales promelas EST project

Unpublished (2005)

Contact: Lindquist, E.A., Richardson, P.

DOE Joint Genome Institute

2800 Mitchell Drive, Walnut Creek, CA 94598, USA

Tel: 925 296 5600

Fax: 925 296 5710

Email: cdna@jgi-psf.org

cDNA Library Preparation: DOE Joint Genome Institute:

http://www.jgi.doe.gov

DNA Sequencing: DOE Joint Genome Institute: http://www.jgi.doe.gov

Naming Conventions: EST name is generated by the concatenation of

the JGI Clone Id and the direction of sequencing. The suffix '.fwd'

indicates a forward sequencing read of the insert. It does not

necessarily reflect the orientation of the insert.

Small Insert: Based upon one or more sequencing reads of this clone

where vector sequence was present at both ends, this clone has been

determined to contain a cDNA insert on the order of 600-1000 bases.

Plate: ANNN 0005 row: j column: 14

High quality sequence stop: 720.

Location/Qualifiers

1..755

/organism="Pimephales promelas"

/mol_type="mRNA"

/db_xref="taxon:90988"

/clone="ANN725"

/tissue_type="Whole"
 /dev_stage="Adult"
 /clone_lib="ANNN Pimephales promelas Whole (L)"
 /notes="Vector: pDONR222; The library was made from dT primed cDNA and cloned into Invitrogen vector pDONR222. Poly A RNA were primed with a Biotin-attB2-Oligo (dT) primer (5'-GGGGCGGCACAACTTTGTACAAGAAAGTTGGGT(T)19-3') and the first strand synthesized using Superscript II (Invitrogen). The second strand was synthesized and an attB1 adaptor was ligated to the 5' end (5'-TCGTCGGGACAACTTTGTACAAGAAAGTTGG-3' and 5'-CCAACCTTTTGTACAAGTTGTCCCC-3'). cDNA was size selected using 1% agarose gel electrophoresis (L -0.5-1.2k, M -1.2-2.5k, H ->2.5k) and then inserted into the vector using site specific recombination (flanking attB sites on cDNA). The work was done at DOE Joint Genome Institute."

ORIGIN

Query Match 76.8%; Score 19.2; DB 8; Length 755;
 Best Local Similarity 87.5%; Pred. No. 9.6e+02;
 Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAACTGGAATCTCAGGCTGAGA 25
 ||||| ||||| ||||| ||||| |||||
 Db 165 AAAAACTGCAATATCAGCCTGAGA 142

RESULT 41
DT107228/c

LOCUS DT107228 755 bp mRNA linear EST 12-AUG-2005
 DEFINITION JGI ANNN8473.fwd ANNN Pimephales promelas Whole (L) Pimephales
 Promelas cDNA clone ANNN8473 5', mRNA sequence.

ACCESSION DT107228
 VERSION DT107228.1 GI:72460123
 KEYWORDS EST.

SOURCE

ORGANISM Pimephales promelas
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
 Cypriniformes; Cyprinidae; Pimephales.

1 (bases 1 to 755)

REFERENCE Richardson, P., Lucas, S., Rokhsar, D., Dettler, J.C., Ng, D.C.,
 Brokstein, P. and Lindquist, E.A.

DOE Joint Genome Institute Pimephales promelas EST project
 Unpublished (2005)

TITLE Contact: Lindquist, E.A., Richardson, P.
 JOURNAL Other ESTs: JGI ANNN8473.rev

COMMENT DOE Joint Genome Institute
 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
 Tel: 925 296 5600
 Fax: 925 296 5710

Email: cdna@jgi-psf.org
 http://www.jgi.doe.gov
 cDNA Library Preparation: DOE Joint Genome Institute:
 DNA Sequencing: DOE Joint Genome Institute: http://www.jgi.doe.gov

Naming Conventions: EST name is generated by the concatenation of the JGI Clone Id and the direction of sequencing. The suffix '.fwd' indicates a forward sequencing read of the insert. It does not necessarily reflect the orientation of the insert.

Small Insert: Based upon one or more sequencing reads of this clone where vector sequence was present at both ends, this clone has been determined to contain a cDNA insert on the order of 600-1000 bases. Plate: ANNN 0089 row: a column: 7
 High quality sequence stop: 752.

FEATURES

source
 1..755
 Location/Qualifiers
 /organism="Pimephales promelas"
 /mol_type="mRNA"
 /db_xref="taxon:90988"
 /clone="ANNN8473"
 /tissue_type="Whole"
 /dev_stage="Adult"
 /clone_lib="ANNN Pimephales promelas Whole (L)"

/note="Vector: pDONR222; The library was made from dT primed cDNA and cloned into Invitrogen vector pDONR222. Poly A RNA were primed with a Biotin-attB2-Oligo (dT) primer (5'-GGGGCGGCACAACTTTGTACAAGAAAGTTGGGT(T)19-3') and the first strand synthesized using Superscript II (Invitrogen). The second strand was synthesized and an attB1 adaptor was ligated to the 5' end (5'-TCGTCGGGACAACTTTGTACAAGAAAGTTGG-3' and 5'-CCAACCTTTTGTACAAGTTGTCCCC-3'). cDNA was size selected using 1% agarose gel electrophoresis (L -0.5-1.2k, M -1.2-2.5k, H ->2.5k) and then inserted into the vector using site specific recombination (flanking attB sites on cDNA). The work was done at DOE Joint Genome Institute."

ORIGIN

Query Match 76.8%; Score 19.2; DB 8; Length 755;
 Best Local Similarity 87.5%; Pred. No. 9.6e+02;
 Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAACTGGAATCTCAGGCTGAGA 25
 ||||| ||||| ||||| ||||| |||||
 Db 167 AAAAACTGCAATATCAGCCTGAGA 144

RESULT 42
DT095581/c

LOCUS DT095581 769 bp mRNA linear EST 12-AUG-2005
 DEFINITION JGI ANNN1867.fwd ANNN Pimephales promelas Whole (L) Pimephales
 Promelas cDNA clone ANNN1867 5', mRNA sequence.

ACCESSION DT095581
 VERSION DT095581.1 GI:72436139
 KEYWORDS EST.

SOURCE

ORGANISM Pimephales promelas
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
 Cypriniformes; Cyprinidae; Pimephales.

1 (bases 1 to 769)

REFERENCE Richardson, P., Lucas, S., Rokhsar, D., Dettler, J.C., Ng, D.C.,
 Brokstein, P. and Lindquist, E.A.

DOE Joint Genome Institute Pimephales promelas EST project
 Unpublished (2005)

TITLE Contact: Lindquist, E.A., Richardson, P.
 JOURNAL Other ESTs: JGI ANNN1867.rev

COMMENT DOE Joint Genome Institute
 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
 Tel: 925 296 5600
 Fax: 925 296 5710

Email: cdna@jgi-psf.org
 http://www.jgi.doe.gov
 cDNA Library Preparation: DOE Joint Genome Institute:
 DNA Sequencing: DOE Joint Genome Institute: http://www.jgi.doe.gov

Naming Conventions: EST name is generated by the concatenation of the JGI Clone Id and the direction of sequencing. The suffix '.fwd' indicates a forward sequencing read of the insert. It does not necessarily reflect the orientation of the insert.

Small Insert: Based upon one or more sequencing reads of this clone where vector sequence was present at both ends, this clone has been determined to contain a cDNA insert on the order of 600-1000 bases. Plate: ANNN 0017 row: f column: 12
 High quality sequence stop: 761.

FEATURES

source
 1..769
 Location/Qualifiers
 /organism="Pimephales promelas"
 /mol_type="mRNA"
 /db_xref="taxon:90988"
 /clone="ANNN1867"
 /tissue_type="Whole"
 /dev_stage="Adult"
 /clone_lib="ANNN Pimephales promelas Whole (L)"
 /note="Vector: pDONR222; The library was made from dT primed cDNA and cloned into Invitrogen vector pDONR222. Poly A RNA were primed with a Biotin-attB2-Oligo (dT)

-1.2-2.5k, H ->2.5k) and then inserted into the vector using site specific recombination (flanking attB sites on cDNA). The work was done at DOE Joint Genome Institute."

ORIGIN

Query Match 76.8%; Score 19.2; DB 8; Length 791;
Best Local Similarity 87.5%; Pred. No. 9.7e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAACTGGAATCTCAGGCTGAGA 25
|||||
DB 167 AAAAACTGCAATATCAGCTGAGA 144
|||||

RESULT 45

BX241295/c 791 bp DNA linear GSS 13-MAR-2003
LOCUS Danio rerio genomic clone DKEY-246B9, genomic survey sequence.
DEFINITION BX241295
ACCESSION BX241295
VERSION BX241295.1 GI:28163629
KEYWORDS GSS.
SOURCE Danio rerio (zebrafish)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.

REFERENCE

1 (bases 1 to 791)
Humphray, S.J., Huckle, E. and Durham, J.L.

AUTHORS

Direct Submission

TITLE

Submitted (13-MAR-2003) The Sanger Institute, Wellcome Trust Genome

JOURNAL

Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquerry@sanger.ac.uk Unpublished

COMMENT

This sequence was generated from the T7 end of BAC 246B9. 246B9 is
part of the Daniokey BAC Library created by R. Plasterk and N.V.

Keygene. Further details:
http://www.sanger.ac.uk/Projects/D_rexio/.

FEATURES

source

1..791
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="DKEY-246B9"
/tissue types="Testis"
/note="Vector pindigoBAC-536"

ORIGIN

Query Match 76.8%; Score 19.2; DB 10; Length 791;
Best Local Similarity 87.5%; Pred. No. 9.7e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAACTGGAATCTCAGGCTGAGA 25
|||||
DB 343 AAAAACTGGAATCCAGCTCTTGA 320
|||||

RESULT 46

BZ452080 824 bp DNA linear GSS 13-DEC-2002
LOCUS BOND84TF BO.1.6.2 KB tot Brassica oleracea genomic clone BOND84,
genomic survey sequence.
DEFINITION BZ452080
ACCESSION BZ452080
VERSION BZ452080.1 GI:26724809
KEYWORDS GSS.

SOURCE

Brassica oleracea
Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.

1 (bases 1 to 824)

REFERENCE

1 (bases 1 to 824)
Ayele, M., Haas, B.J., Kumar, N., Wu, H., Xiao, Y., Van Aken, S.,
Utterback, T.R., Wortman, J.R., White, O.R. and Town, C.D.

AUTHORS

Whole genome shotgun sequencing of Brassica oleracea and its
application to gene discovery and annotation in Arabidopsis

JOURNAL
PUBMED
COMMENT

Genome Res. 15 (4), 487-495 (2005)
15805490
Other GSSs: BOND84TR
Contact: Chris Town

TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org

DNA is from a doubled haploid provided by Tom Osborn.

Seq primer: TF

Class: Sheared ends.

FEATURES

source

Location/Qualifiers
1..824
/organism="Brassica oleracea"
/mol_type="genomic DNA"
/strain="T01000DH3"
/db_xref="taxon:3712"
/clone="BOND84"
/clone_lib="BO.1.6.2_KB_tot"
/note="Vector: pHOS1; Site 1: BatXI; 1.6-2 kb sheared
total DNA inserted into pHOS1 using BatXI linkers"

ORIGIN

Query Match 76.8%; Score 19.2; DB 9; Length 824;
Best Local Similarity 87.5%; Pred. No. 9.8e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAACTGGAATCTCAGGCTGAGA 25
|||||

DB 524 AAAAACTGGTAATCAGGCTGAGA 547
|||||

RESULT 47

BUI25751 861 bp mRNA linear EST 25-NOV-2002
LOCUS 603149763F1 CSEQCHL19 Gallus gallus cDNA clone CHEST154n16 5', mRNA
DEFINITION BUI25751
ACCESSION BUI25751
VERSION BUI25751.1 GI:25336704
KEYWORDS EST.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.

1 (bases 1 to 861)
Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E.,
Fong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.
A Comprehensive Collection of Chicken cDNAs
Curr. Biol. 12 (22), 1965-1969 (2002)
12445392

COMMENT

Contact: Simon Hubbard
Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology
(UMIST)

PO Box 88, Manchester, M60 1QD, UK

Tel: 01612008930

Fax: 01612360409

Email: Simon.Hubbard@umist.ac.uk.

FEATURES

source

Location/Qualifiers
1..861
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="Compton Line 151"
/db_xref="taxon:9031"
/clone="CHEST154n16"
/sex="Female"
/dev stage="adult"
/lab_host="DH10B"
/clone_lib="CSEQCHL19"

/note="Organ: liver; Vector: pBluescript II KS(+); Site 1:
EcoRI; Site 2: NotI; Modification of pBluescript II KS(+)

[Stratagene] vector to accommodate cDNA produced with the T-trimmed protocol (Construction of uni-directionally cloned cDNA libraries from messenger RNA for improved 3' end DNA sequencing by Glenn Fu, et al. U.S. Patent # 6,387,624). Cut pBluescript II KS(+) with NotI and EcoRI. Ligate in double stranded adaptor containing BsgI and BamHI sites [5'ggcgcgtgagcccgagtcgcaaaaaag] [5'aattcttttttcggatccggggtcgacgc]"

ORIGIN

Query Match 76.8%; Score 19.2; DB 5; Length 861;
Best Local Similarity 87.5%; Pred. No. 9.9e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAAGTGAATCTCAGGCTGAGA 25

DB 829 AAAAAGTGAATCACAGTCTGAAA 852

RESULT 48

CD126154 508 bp mRNA linear EST 14-SEP-2003
LOCUS MF1-0002U-V053-E04-U.B MF1-0002 Schistosoma mansoni cDNA clone
DEFINITION MF1-0002U-V053-E04.B, mRNA sequence.
ACCESSION CD126154
VERSION CD126154.1 GI:34664163
KEYWORDS EST.
SOURCE Schistosoma mansoni
ORGANISM Schistosoma mansoni

REFERENCE

1 (bases 1 to 508)
Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea;
Strigeidida; Schistosomatoidea; Schistosomatidae; Schistosoma.

AUTHORS

Verjovski-Almeida, S., DeMarco, R., Martins, E.A.L., Guimaraes, P.E.M., Ojopi, E.P.B., Paquola, A.C.M., Piazza, J.P., Nishiyama, M.Y. Jr., Kitajima, J.P., Adams, R.E., Ashton, P.D., Bonaldo, M.F., Coulson, P.S., Dillon, G.P., Farias, L.P., Gregorio, S.P., Ho, P.L., Leite, R.A., Malaquias, L.C.C., Marques, R.C.P., Miyasato, P.A., Nascimento, A.L.T.O., Ohlweiler, F.P., Reis, E.M., Ribeiro, M.A., Sa, R.G., Stukart, G.C., Soares, M.B., Gargioni, C., Kawano, T., Rodrigues, V., Madeira, A.M.B.N., Wilson, R.A., Menck, C.F.M., Setubal, J.C., Leite, L.C.C. and Dias-Neto, E.

Transcriptome analysis of the acelomate human parasite Schistosoma mansoni

JOURNAL

Nat. Genet. 35 (2), 148-157 (2003)

PUBMED

12973350

COMMENT

Contact: Dr. Sergio Verjovski-Almeida
Departamento de Bioquímica
Instituto de Química - Universidade de São Paulo
Av. Prof. Lineu Prestes 748 sala 1200, 05508-900 São Paulo - SP, Brazil
Tel: +55-11-3091-2173
Fax: +55-11-3091-2186

Email: verjo@iq.usp.br

This sequence was derived from the FAPESP Schistosoma mansoni EST Genome Project. All sequences in the project were assembled and annotated. This entry and all the assembled sequences can be seen in the following URL <http://bioinfo.iq.usp.br/schisto/>
Plate: MF1-0002U-V053 row: 4 column: E.

FEATURES

source

1..508
/organism="Schistosoma mansoni"
/mol_type="mRNA"
/db_xref="taxon:6183"
/clone="MF1-0002U-V053-E04.B"
/sex="female"
/dev_stage="adult"
/lab_host="Mesocricetus auratus"
/clone_lib="MF1-0002"
/note="Vector: SureClone"

ORIGIN

Query Match 76.0%; Score 19; DB 6; Length 508;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 AAAAAGTGAATCTCAGGC 20
DB 481 AAAAAGTGAATCTCAGGC 499

RESULT 49

BU254969 634 bp mRNA linear EST 26-NOV-2002
LOCUS 603414046F1 CSEQCHN38 Gallus gallus cDNA clone CHEST335111 5', mRNA
DEFINITION sequence.
ACCESSION BU254969
VERSION BU254969.1 GI:25511230
KEYWORDS EST.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus

REFERENCE 1 (bases 1 to 634)
AUTHORS Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E., Fong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.
TITLE A Comprehensive Collection of Chicken cDNAs
JOURNAL Curr. Biol. 12 (22), 1965-1969 (2002)
PUBMED 12445392
COMMENT Contact: Simon Hubbard
Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology (UMIST)
PO Box 88, Manchester, M60 1QD, UK
Tel: 01612008930
Fax: 01612360409
Email: Simon.Hubbard@umist.ac.uk.

FEATURES
source
1..634
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="White Leghorn, Hisex"
/db_xref="taxon:9031"
/clone="CHEST335111"
/dev_stage="36"
/lab_host="DH10B"
/clone_lib="CSEQCHN38"
/note="Organ: limbs; Vector: pBluescript II KS(+); Site: 1: EcoRI; Site 2: NotI; This normalized library was constructed from 1 million independent clones. cDNA synthesis was initiated using an oligo(dT) primer, using methylated C in the first strand synthesis reaction. Following this first strand reaction, double-stranded cDNA was blunted, ligated to NotI adapters, digested with EcoRI, size-selected, and cloned into the NotI and EcoRI compatible sites of a custom modified MCS of the pBluescript (KS+) vector. The library was normalized in 2 rounds using conditions adapted from Soares et al., PNAS (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6 (1996): 791, except that a significantly longer reannealing hybridization was used."

ORIGIN

Query Match 76.0%; Score 19; DB 5; Length 634;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 AAACTGGAATCTCAGGCTG 22

DB 490 AAACTGGAATCTCAGGCTG 508

RESULT 50

BU338343 659 bp mRNA linear EST 28-NOV-2002
LOCUS BU338343
DEFINITION 603515585F1 CSEQCHN66 Gallus gallus cDNA clone CHEST454j22 5', mRNA

sequence.
ACCESSION BU338343
VERSION BU338343.1 GI:25846344
KEYWORDS EST.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 659)
AUTHORS Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,
Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.
TITLE A Comprehensive Collection of Chicken cDNAs
JOURNAL Curr. Biol. 12 (22), 1965-1969 (2002)
PUBMED 12445392
COMMENT Contact: Simon Hubbard
Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology
(UMIST)
PO Box 88, Manchester, M60 1QD, UK
Tel: 01612008930
Fax: 01612360409
Email: Simon.Hubbard@umist.ac.uk.
Location/Qualifiers
FEATURES
source
1..659
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="White Leghorn, Hisex"
/db_xref="taxon:9031"
/clone="CHEST454j22"
/dev_stage="16 day embryo"
/lab_host="DH10B"
/clone_lib="CSEQCHN66"
/note="Organ: brain; Vector: pBluescript II KS(+); Site_1:
EcoRI; Site_2: NotI; This normalized library was
constructed from 1 million independent clones. cDNA
synthesis was initiated using an oligo(dT) primer, using
methylated C in the first strand synthesis reaction.
Following this first strand reaction, double-stranded cDNA
was bluntended, ligated to NotI adapters, digested with
EcoRI, size-selected, and cloned into the NotI and EcoRI
compatible sites of a custom modified MCS of the
pBluescript (KS+) vector. The library was normalized in 2
rounds using conditions adapted from Soares et al., PNAS
(1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
(1996): 791, except that a significantly longer
reannealing hybridization was used."

ORIGIN

Query Match 76.0%; Score 19; DB 5; Length 659;
Best Local Similarity 100.0%; Pred.No. 1.2e+03;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 AACTGGAAATCTCAGGCTG 22
|||||
Db 281 AACTGGAAATCTCAGGCTG 299

Search completed: February 3, 2006, 22:02:39
Job time : 2959.67 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 3, 2006, 11:35:31 ; Search time 80.6667 Seconds
(without alignments)
550.897 Million cell updates/sec

Title: US-10-719-900-30

Perfect score: 25

Sequence: 1 aaaaactggaatcaggtgaga 25

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database :

Issued Patents NA:*

- 1: /cgn2_6/ptodata/1/ina/1 COMB.seq:*
- 2: /cgn2_6/ptodata/1/ina/5 COMB.seq:*
- 3: /cgn2_6/ptodata/1/ina/6A COMB.seq:*
- 4: /cgn2_6/ptodata/1/ina/6B COMB.seq:*
- 5: /cgn2_6/ptodata/1/ina/H COMB.seq:*
- 6: /cgn2_6/ptodata/1/ina/PTUS COMB.seq:*
- 7: /cgn2_6/ptodata/1/ina/PP COMB.seq:*
- 8: /cgn2_6/ptodata/1/ina/RE COMB.seq:*
- 9: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	19.4	77.6	39552	3	US-09-949-016-11799
2	19.2	76.8	11003	3	US-09-949-016-13166
3	18.8	75.2	51955	3	US-09-949-016-15854
4	18.6	74.4	192506	3	US-09-949-016-15830
5	18.2	72.8	95318	3	US-09-949-016-11784
6	18.2	72.8	95318	3	US-09-949-016-13998
7	17.8	71.2	601	3	US-09-949-016-45703
8	17.8	71.2	150394	3	US-09-949-016-13042
9	17.6	70.4	601	3	US-09-949-016-48777
10	17.6	70.4	31318	3	US-09-949-016-12495
11	17.6	70.4	31319	3	US-09-949-016-15963
12	17.6	70.4	51928	3	US-09-949-016-13184
13	17.2	68.8	9300	3	US-09-949-016-12482
14	17.2	68.8	9301	3	US-09-949-016-12888
15	17.2	68.8	26619	3	US-09-949-016-15030
16	17.2	68.8	41612	3	US-09-949-016-12769
17	17.2	68.8	41639	3	US-09-949-016-12773
18	17.2	68.8	77388	3	US-09-949-016-13496
19	17.2	68.8	112705	3	US-09-949-016-15630
20	17.2	68.8	115388	3	US-09-949-016-14981
21	17.2	68.8	127771	3	US-09-949-016-14982
22	17.2	68.8	264206	3	US-09-949-016-12731
23	17.2	68.8	264304	3	US-09-949-016-13249
24	17	68.0	234	3	US-09-248-796A-7364

25	17	68.0	304	3	US-09-513-999C-24054	Sequence 24054, A
26	17	68.0	601	3	US-09-949-016-145029	Sequence 145029, A
27	17	68.0	601	3	US-09-949-016-145030	Sequence 145030, A
28	17	68.0	601	3	US-09-949-016-170085	Sequence 170085, A
29	17	68.0	631	3	US-09-513-999C-14974	Sequence 14974, A
30	17	68.0	1111	3	US-09-167-206-9	Sequence 9, Appli
31	17	68.0	1138	3	US-10-164-595-33	Sequence 3, Appli
32	17	68.0	1424	3	US-08-956-171B-566	Sequence 566, App
33	17	68.0	1424	3	US-08-781-986A-566	Sequence 566, App
34	17	68.0	1717	3	US-08-956-171B-481	Sequence 481, App
35	17	68.0	1717	3	US-08-781-986A-481	Sequence 481, App
36	17	68.0	1858	3	US-09-347-819-1	Sequence 1, Appli
37	17	68.0	2880	3	US-10-104-047-859	Sequence 859, App
38	17	68.0	7736	3	US-09-949-016-17579	Sequence 17579, A
39	17	68.0	20347	3	US-09-949-016-16752	Sequence 16752, A
40	17	68.0	22965	3	US-09-949-016-15862	Sequence 15862, A
41	17	68.0	23533	3	US-09-949-016-15377	Sequence 15377, A
42	17	68.0	42989	3	US-09-949-016-12291	Sequence 12291, A
43	17	68.0	42992	3	US-09-949-016-15428	Sequence 15428, A
44	17	68.0	49677	3	US-09-949-002-817	Sequence 817, App
45	17	68.0	49818	3	US-09-949-016-12458	Sequence 12458, A
46	17	68.0	49829	3	US-09-949-016-14011	Sequence 14011, A
47	17	68.0	85875	3	US-09-949-016-12333	Sequence 12333, A
48	17	68.0	85875	3	US-09-949-016-15956	Sequence 15956, A
49	17	68.0	96327	3	US-09-949-016-16541	Sequence 16541, A
50	17	68.0	105189	3	US-09-949-016-13029	Sequence 13029, A
51	17	68.0	129327	3	US-09-949-016-12257	Sequence 12257, A
52	17	68.0	129327	3	US-09-949-016-15368	Sequence 15368, A
53	17	68.0	177251	3	US-09-949-016-15941	Sequence 15941, A
54	17	68.0	194790	3	US-09-949-016-15393	Sequence 15393, A
55	17	68.0	260247	3	US-09-949-016-13358	Sequence 13358, A
56	17	68.0	1230025	3	US-09-198-452A-1	Sequence 1, Appli
57	17	68.0	1230230	3	US-09-438-185A-1	Sequence 1, Appli
58	16.8	67.2	133	3	US-09-513-999C-27290	Sequence 27290, A
59	16.8	67.2	601	3	US-09-949-016-59983	Sequence 59983, A
60	16.8	67.2	601	3	US-09-949-016-117283	Sequence 117283, A
61	16.8	67.2	601	3	US-09-949-016-117352	Sequence 117352, A
62	16.8	67.2	31214	3	US-09-949-016-17097	Sequence 17097, A
63	16.8	67.2	31214	3	US-09-949-016-17098	Sequence 17098, A
64	16.8	67.2	32958	3	US-09-949-016-15677	Sequence 15677, A
65	16.8	67.2	36242	3	US-09-949-016-12996	Sequence 12996, A
66	16.8	67.2	36242	3	US-09-949-016-12997	Sequence 12997, A
67	16.8	67.2	36242	3	US-09-949-016-12998	Sequence 12998, A
68	16.8	67.2	36242	3	US-09-949-016-12999	Sequence 12999, A
69	16.8	67.2	36242	3	US-09-949-016-13000	Sequence 13000, A
70	16.8	67.2	46626	3	US-09-949-016-13390	Sequence 13390, A
71	16.8	67.2	60595	3	US-09-949-002-650	Sequence 650, App
72	16.8	67.2	80595	3	US-09-949-002-706	Sequence 706, App
73	16.8	67.2	83450	3	US-09-811-469-3	Sequence 3, Appli
74	16.8	67.2	83450	3	US-10-370-659-3	Sequence 3, Appli
75	16.8	67.2	168394	3	US-09-949-016-13002	Sequence 13002, A
76	16.6	66.4	217	3	US-09-513-999C-3132	Sequence 3132, Ap
77	16.6	66.4	472	3	US-09-270-767-6329	Sequence 6329, Ap
78	16.6	66.4	472	3	US-09-270-767-6329	Sequence 21611, A
79	16.6	66.4	601	3	US-09-270-767-21611	Sequence 53654, A
80	16.6	66.4	601	3	US-09-949-016-53654	Sequence 80518, A
81	16.6	66.4	651	3	US-09-270-767-7426	Sequence 7426, Ap
82	16.6	66.4	651	3	US-09-270-767-22708	Sequence 22708, A
83	16.6	66.4	858	3	US-09-621-976-3699	Sequence 3699, Ap
84	16.6	66.4	885	3	US-09-270-767-8281	Sequence 8281, Ap
85	16.6	66.4	885	3	US-09-270-767-23563	Sequence 23563, A
86	16.6	66.4	1308	3	US-09-248-796A-2964	Sequence 2964, App
87	16.6	66.4	1990	3	US-09-149-476-281	Sequence 281, App
88	16.6	66.4	2041	3	US-09-149-476-131	Sequence 131, App
89	16.6	66.4	2334	2	US-08-062-632-4	Sequence 4, Appli
90	16.6	66.4	8180	3	US-09-949-016-14422	Sequence 14422, A
91	16.6	66.4	26173	3	US-09-453-702B-69	Sequence 69, Appli
92	16.6	66.4	26173	3	US-10-114-170-69	Sequence 12485, A
93	16.6	66.4	34011	3	US-09-949-016-12485	Sequence 12485, A
94	16.6	66.4	36618	3	US-09-949-016-12961	Sequence 12961, A
95	16.6	66.4	48908	3	US-09-453-702B-137	Sequence 137, App
96	16.6	66.4	48908	3	US-10-114-170-137	Sequence 137, App
97	16.6	66.4	60137	3	US-09-949-016-14735	Sequence 14735, A

```

98 16.6 66.4 60137 3 US-09-949-016-14912 Sequence 14912, A
c 99 16.6 66.4 72549 3 US-09-949-016-16477 Sequence 16477, A
100 16.6 66.4 88557 3 US-09-949-016-17028 Sequence 17028, A
101 16.6 66.4 108169 3 US-09-949-016-12898 Sequence 12898, A
102 16.6 66.4 108169 3 US-09-949-016-12898 Sequence 12898, A
103 16.6 66.4 108169 3 US-09-949-016-12898 Sequence 12898, A
104 16.6 66.4 109025 3 US-09-949-016-12609 Sequence 12609, A
105 16.6 66.4 109025 3 US-09-949-016-12609 Sequence 12609, A
106 16.6 66.4 109025 3 US-09-949-016-12609 Sequence 12609, A
107 16.6 66.4 134434 3 US-09-949-016-17362 Sequence 17362, A
108 16.6 66.4 134434 3 US-09-949-016-17362 Sequence 17362, A
109 16.6 66.4 134434 3 US-09-949-016-17362 Sequence 17362, A
110 16.6 66.4 134434 3 US-09-949-016-17362 Sequence 17362, A
111 16.6 66.4 134434 3 US-09-949-016-17362 Sequence 17362, A
112 16.6 66.4 134434 3 US-09-949-016-17362 Sequence 17362, A
113 16.6 66.4 134434 3 US-09-949-016-17362 Sequence 17362, A
114 16.6 66.4 134434 3 US-09-949-016-17362 Sequence 17362, A
115 16.6 66.4 134434 3 US-09-949-016-17362 Sequence 17362, A
116 16.6 66.4 134434 3 US-09-949-016-17362 Sequence 17362, A
117 16.6 66.4 134434 3 US-09-949-016-17362 Sequence 17362, A
118 16.6 66.4 134434 3 US-09-949-016-17362 Sequence 17362, A
119 16.6 66.4 134434 3 US-09-949-016-17362 Sequence 17362, A
120 16.6 66.4 134434 3 US-09-949-016-17362 Sequence 17362, A
121 16.6 66.4 134434 3 US-09-949-016-17362 Sequence 17362, A
122 16.6 66.4 134434 3 US-09-949-016-17362 Sequence 17362, A
123 16.6 66.4 134434 3 US-09-949-016-17362 Sequence 17362, A
124 16.6 66.4 134434 3 US-09-949-016-17362 Sequence 17362, A
125 16.6 66.4 134434 3 US-09-949-016-17362 Sequence 17362, A
126 16.6 66.4 134434 3 US-09-949-016-17362 Sequence 17362, A
127 16.6 66.4 134434 3 US-09-949-016-17362 Sequence 17362, A
128 16.6 66.4 134434 3 US-09-949-016-17362 Sequence 17362, A
129 16.6 66.4 134434 3 US-09-949-016-17362 Sequence 17362, A
130 16.6 66.4 134434 3 US-09-949-016-17362 Sequence 17362, A
131 16.6 66.4 134434 3 US-09-949-016-17362 Sequence 17362, A
132 16.6 66.4 134434 3 US-09-949-016-17362 Sequence 17362, A
133 16.6 66.4 134434 3 US-09-949-016-17362 Sequence 17362, A
134 16.6 66.4 134434 3 US-09-949-016-17362 Sequence 17362, A
135 16.6 66.4 134434 3 US-09-949-016-17362 Sequence 17362, A
136 16.6 66.4 134434 3 US-09-949-016-17362 Sequence 17362, A
137 16.6 66.4 134434 3 US-09-949-016-17362 Sequence 17362, A
138 16.6 66.4 134434 3 US-09-949-016-17362 Sequence 17362, A
139 16.6 66.4 134434 3 US-09-949-016-17362 Sequence 17362, A
140 16.6 66.4 134434 3 US-09-949-016-17362 Sequence 17362, A
141 16.6 66.4 134434 3 US-09-949-016-17362 Sequence 17362, A
142 16.6 66.4 134434 3 US-09-949-016-17362 Sequence 17362, A
143 16.6 66.4 134434 3 US-09-949-016-17362 Sequence 17362, A
144 16.6 66.4 134434 3 US-09-949-016-17362 Sequence 17362, A
145 16.6 66.4 134434 3 US-09-949-016-17362 Sequence 17362, A
146 16.6 66.4 134434 3 US-09-949-016-17362 Sequence 17362, A
147 16.6 66.4 134434 3 US-09-949-016-17362 Sequence 17362, A
148 16.6 66.4 134434 3 US-09-949-016-17362 Sequence 17362, A
149 16.6 66.4 134434 3 US-09-949-016-17362 Sequence 17362, A
150 16.6 66.4 134434 3 US-09-949-016-17362 Sequence 17362, A
```

ALIGNMENTS

```

RESULT 1
US-09-949-016-11799
; Sequence 11799, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498

RESULT 3
US-09-949-016-15854/c
; Sequence 15854, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498

Query Match 76.8%; Score 19.2; DB 3; Length 11003;
Best Local Similarity 87.5%; Pred. No. 28;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAACTGGAATCTCAGGCTGAGA 25
DB 7198 AGAACTGGAATCTCAGGCTGAGA 7221

RESULT 3
US-09-949-016-15854/c
; Sequence 15854, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
```

; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15854
; LENGTH: 55195
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15854

Query Match 75.2%; Score 18.8; DB 3; Length 55195;
Best Local Similarity 90.9%; Pred. No. 56;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 AAAAAGCTTGAATCTCAGGCTGAG 24
|||||
Db 40869 ACAACTGGAATCTCAGCTGAG 40848

RESULT 4
US-09-949-016-15830/c
; Sequence 15830, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15830
; LENGTH: 192506
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15830

Query Match 74.4%; Score 18.6; DB 3; Length 192506;
Best Local Similarity 84.0%; Pred. No. 85;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAGCTTGAATCTCAGGCTGAGA 25
|||||
Db 169823 AGAAGCTTGAATCTCAGGCTGAGA 169799

RESULT 5
US-09-949-016-11784/c
; Sequence 11784, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11784
; LENGTH: 95318

; TYPE: DNA
; ORGANISM: Human
US-09-949-016-11784

Query Match 72.8%; Score 18.2; DB 3; Length 95318;
Best Local Similarity 87.0%; Pred. No. 1.2e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 AAAAAGCTTGAATCTCAGGCTGAGA 25
|||||
Db 45815 AAAATTGAATCTCAGCTGAGA 45793

RESULT 6
US-09-949-016-13998/c
; Sequence 13998, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13998
; LENGTH: 95318
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13998

Query Match 72.8%; Score 18.2; DB 3; Length 95318;
Best Local Similarity 87.0%; Pred. No. 1.2e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 AAAAAGCTTGAATCTCAGGCTGAGA 25
|||||
Db 45815 AAAATTGAATCTCAGCTGAGA 45793

RESULT 7
US-09-949-016-45703
; Sequence 45703, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 45703
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-45703

Query Match 71.2%; Score 17.8; DB 3; Length 601;

```
Best Local Similarity 90.5%; Pred. No. 82;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 AAAAAGTGAATCTCAGGCTGA 23
    ||||| ||||| ||||| |||||
Db 46 AAAAAGTGAATCTCAGGCTGA 66

RESULT 8
US-09-949-016-13042/c
; Sequence 13042, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13042
; LENGTH: 150394
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(150394)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13042

Query Match 71.2%; Score 17.8; DB 3; Length 150394;
Best Local Similarity 90.5%; Pred. No. 1.9e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 AAAAAGTGAATCTCAGGCTGA 23
    ||||| ||||| ||||| |||||
Db 124943 AAAAAGTGAATCTCAGGCTGA 124923

RESULT 9
US-09-949-016-48777/c
; Sequence 48777, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 48777
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-48777

Query Match 70.4%; Score 17.6; DB 3; Length 601;
Best Local Similarity 83.3%; Pred. No. 1e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AAAAAGTGAATCTCAGGCTGA 25
    ||||| ||||| ||||| |||||
Db 249 AAAAAGGACACTCAGGCTGA 226

RESULT 10
US-09-949-016-12495
; Sequence 12495, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12495
; LENGTH: 31318
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12495

Query Match 70.4%; Score 17.6; DB 3; Length 31318;
Best Local Similarity 83.3%; Pred. No. 1.9e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAGTGAATCTCAGGCTGA 24
    ||||| ||||| ||||| |||||
Db 11006 AAAAAGTGAATCTCAGGCTGA 11029

RESULT 11
US-09-949-016-15963
; Sequence 15963, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15963
; LENGTH: 31319
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15963

Query Match 70.4%; Score 17.6; DB 3; Length 31319;
Best Local Similarity 83.3%; Pred. No. 1.9e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAGTGAATCTCAGGCTGA 24
    ||||| ||||| ||||| |||||
Db 11006 AAAAAGTGAATCTCAGGCTGA 11029
```

RESULT 16
US-09-949-016-12769
; Sequence 12769, Application US/09949016
; Patent No. 6812339

; Sequence 14981, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14981
; LENGTH: 115388
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-14981

Query Match 68.8%; Score 17.2; DB 3; Length 115388;
Best Local Similarity 86.4%; Pred. No. 3.6e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAACTGGAATCTCAGGCTG 22
Db 74064 AAAGAACTGGAATTTTCAGGATG 74085

RESULT 21
US-09-949-016-14982
; Sequence 14982, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14982
; LENGTH: 127771
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-14982

Query Match 68.8%; Score 17.2; DB 3; Length 127771;
Best Local Similarity 86.4%; Pred. No. 3.6e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAACTGGAATCTCAGGCTG 22
Db 74064 AAAGAACTGGAATTTTCAGGATG 74085

RESULT 22
US-09-949-016-12731/c
; Sequence 12731, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12731
; LENGTH: 264206
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12731

Query Match 68.8%; Score 17.2; DB 3; Length 264206;
Best Local Similarity 86.4%; Pred. No. 4e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 AAACCTGGAATCTCAGGCTGAGA 25
Db 245257 AAACCTGGCATCTCTGGCTAAGA 245236

RESULT 23
US-09-949-016-13249/c
; Sequence 13249, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13249
; LENGTH: 264304
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13249

Query Match 68.8%; Score 17.2; DB 3; Length 264304;
Best Local Similarity 86.4%; Pred. No. 4e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 AAACCTGGAATCTCAGGCTGAGA 25
Db 245257 AAACCTGGCATCTCTGGCTAAGA 245236

RESULT 24
US-09-248-796A-7364
; Sequence 7364, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICA
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725


```
RESULT 28
US-09-949-016-170085/c
; Sequence 170085, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 170085
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-170085

Query Match      68.0%; Score 17; DB 3; Length 601;
Best Local Similarity 80.0%; Pred. No. 1.9e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 AAAAACTGGAATCTCAGGCTGAGA 25
Db 182 AAAAATGTGAAACTGAGGCTTAGA 158

RESULT 29
US-09-513-999C-14974
; Sequence 14974, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 14974
; LENGTH: 631
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 137
; OTHER INFORMATION: r=a or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 138
; OTHER INFORMATION: r=a or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 182
; OTHER INFORMATION: n=a, g, c or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 201
; OTHER INFORMATION: n=a, g, c or t
; FEATURE:
```

```
; NAME/KEY: misc_feature
; LOCATION: 356
; OTHER INFORMATION: m=a or c
US-09-513-999C-14974

Query Match      68.0%; Score 17; DB 3; Length 631;
Best Local Similarity 80.0%; Pred. No. 2e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 AAAAACTGGAATCTCAGGCTGAGA 25
Db 516 AAGTCACTGGAGGCTCAGGCTGAGA 540

RESULT 30
US-09-167-206-9
; Sequence 9, Application US/09167206A
; Patent No. 6476193
; GENERAL INFORMATION:
; APPLICANT: Nandabalan, Krishnan
; APPLICANT: Schulz, Vincent P.
; APPLICANT: Yang, Meijia
; TITLE OF INVENTION: Nlki PROTEIN AND Nlki PROTEIN COMPLEXES
; FILE REFERENCE: 15966-521 Nlki protein complexes
; CURRENT APPLICATION NUMBER: US/09/167,206A
; CURRENT FILING DATE: 1998-10-06
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 111
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (57)..(908)
US-09-167-206-9

Query Match      68.0%; Score 17; DB 3; Length 111;
Best Local Similarity 80.0%; Pred. No. 2.1e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 AAAAACTGGAATCTCAGGCTGAGA 25
Db 669 AAGTCACTGGAGGCTCAGGCTGAGA 693

RESULT 31
US-10-164-595-33
; Sequence 33, Application US/10164595
; Patent No. 6657054
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies, Inc
; TITLE OF INVENTION: Regulated Angiogenesis Genes and Polypeptides
; FILE REFERENCE: IU 103 R1
; CURRENT APPLICATION NUMBER: US/10/164,595
; CURRENT FILING DATE: 2002-06-10
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 33
; LENGTH: 1138
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (231)..(965)
; OTHER INFORMATION:
US-10-164-595-33

Query Match      68.0%; Score 17; DB 3; Length 1138;
Best Local Similarity 80.0%; Pred. No. 2.2e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 AAAAACTGGAATCTCAGGCTGAGA 25
```

```

; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,986A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Benson, Bob
; REGISTRATION NUMBER: 30,446
; REFERENCE/DOCKET NUMBER: PB248PP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 566:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1424 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-08-781-986A-566

Query Match      68.0%; Score 17; DB 3; Length 1424;
Best Local Similarity 80.0%; Pred. No. 2.2e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 AAAAACTGGAATCTCAGGCTGAGA 25
Db 385 AACAAATCGGAATCTAATGCTGAGA 361

RESULT 34
US-08-956-171E-481
; Sequence 481, Application US/08956171E
; Patent No. 6593114
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; Gil H. Choi
; Patrick S. Dillon
; Craig A. Rosen
; Steven C. Barash
; Michael R. Fannon
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5256
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/956,171E
; FILING DATE: 20-Oct-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/009,861
; FILING DATE: January 5, 1996
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (240) 314-1224
; TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 566:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1424 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 566:

US-08-956-171E-566
Query Match      68.0%; Score 17; DB 3; Length 1424;
Best Local Similarity 80.0%; Pred. No. 2.2e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 AAAAACTGGAATCTCAGGCTGAGA 25
Db 385 AACAAATCGGAATCTAATGCTGAGA 361

RESULT 33
US-08-781-986A-566/c
; Sequence 566, Application US/08781986A
; Patent No. 6737248
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; Gil H. Choi
; Patrick S. Dillon
; Craig A. Rosen
; Steven C. Barash
; Michael R. Fannon
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,986A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Benson, Bob
; REGISTRATION NUMBER: 30,446
; REFERENCE/DOCKET NUMBER: PB248PP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 566:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1424 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 566:

US-08-781-986A-566/c
Query Match      68.0%; Score 17; DB 3; Length 1424;
Best Local Similarity 80.0%; Pred. No. 2.2e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 AAAAACTGGAATCTCAGGCTGAGA 25
Db 385 AACAAATCGGAATCTAATGCTGAGA 361

RESULT 32
US-08-956-171E-566/c
; Sequence 566, Application US/08956171E
; Patent No. 6593114
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; Gil H. Choi
; Patrick S. Dillon
; Craig A. Rosen
; Steven C. Barash
; Michael R. Fannon
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5256
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/956,171E
; FILING DATE: 20-Oct-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/009,861
; FILING DATE: January 5, 1996
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (240) 314-1224
; TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 566:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1424 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 566:

US-08-956-171E-566
Query Match      68.0%; Score 17; DB 3; Length 1424;
Best Local Similarity 80.0%; Pred. No. 2.2e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 AAAAACTGGAATCTCAGGCTGAGA 25
Db 385 AACAAATCGGAATCTAATGCTGAGA 361

RESULT 31
US-08-781-986A-566/c
; Sequence 566, Application US/08781986A
; Patent No. 6737248
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; Gil H. Choi
; Patrick S. Dillon
; Craig A. Rosen
; Steven C. Barash
; Michael R. Fannon
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,986A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Benson, Bob
; REGISTRATION NUMBER: 30,446
; REFERENCE/DOCKET NUMBER: PB248PP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 566:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1424 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 566:

US-08-781-986A-566/c
Query Match      68.0%; Score 17; DB 3; Length 1424;
Best Local Similarity 80.0%; Pred. No. 2.2e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 AAAAACTGGAATCTCAGGCTGAGA 25
Db 385 AACAAATCGGAATCTAATGCTGAGA 361

RESULT 30
US-08-956-171E-481
; Sequence 481, Application US/08956171E
; Patent No. 6593114
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; Gil H. Choi
; Patrick S. Dillon
; Craig A. Rosen
; Steven C. Barash
; Michael R. Fannon
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5256
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/956,171E
; FILING DATE: 20-Oct-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/009,861
; FILING DATE: January 5, 1996
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (240) 314-1224
; TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 566:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1424 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 566:

US-08-956-171E-566
Query Match      68.0%; Score 17; DB 3; Length 1424;
Best Local Similarity 80.0%; Pred. No. 2.2e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 AAAAACTGGAATCTCAGGCTGAGA 25
Db 385 AACAAATCGGAATCTAATGCTGAGA 361

RESULT 29
US-08-781-986A-566/c
; Sequence 566, Application US/08781986A
; Patent No. 6737248
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; Gil H. Choi
; Patrick S. Dillon
; Craig A. Rosen
; Steven C. Barash
; Michael R. Fannon
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,986A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Benson, Bob
; REGISTRATION NUMBER: 30,446
; REFERENCE/DOCKET NUMBER: PB248PP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 566:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1424 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 566:

US-08-781-986A-566/c
Query Match      68.0%; Score 17; DB 3; Length 1424;
Best Local Similarity 80.0%; Pred. No. 2.2e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 AAAAACTGGAATCTCAGGCTGAGA 25
Db 385 AACAAATCGGAATCTA
```

APPLICATION NUMBER: 08/781,986
FILING DATE: January 3, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Mark J. Hyman
REGISTRATION NUMBER: 46,789
REFERENCE/DOCKET NUMBER: PB248P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (240) 314-1224
TELEFAX: (301) 309-8439
INFORMATION FOR SEQ ID NO: 481:
SEQUENCE CHARACTERISTICS:
LENGTH: 1717 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 481:
US-08-956-171E-481

Query Match 68.0%; Score 17; DB 3; Length 1717;
Best Local Similarity 80.0%; Pred. No. 2.3e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAACTGGATCTCAGGCTGAGA 25
||||| ||||| ||||| ||||| |||||
Db 817 AAAAACTGGATCTCAGGCTTAGA 841

RESULT 35
US-08-781-986A-481
; Sequence 481, Application US/08781986A
; Patent No. 6737248
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,986A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Benson, Bob
; REGISTRATION NUMBER: 30,446
; REFERENCE/DOCKET NUMBER: PB248PP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 481:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1717 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-08-781-986A-481

Query Match 68.0%; Score 17; DB 3; Length 1717;
Best Local Similarity 80.0%; Pred. No. 2.3e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAACTGGATCTCAGGCTGAGA 25
||||| ||||| ||||| ||||| |||||
Db 817 AAAAACTGGATCTCAGGCTTAGA 841

RESULT 36
US-09-347-819-1
; Sequence 1, Application US/09347819
; Patent No. 6184036
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Rafalski, J. Antoni
; APPLICANT: Falco, S. Carl
; TITLE OF INVENTION: Ornithine Biosynthesis Enzymes
; FILE REFERENCE: BB-1174-C
; CURRENT APPLICATION NUMBER: US/09/347,819
; CURRENT FILING DATE: 1999-07-02
; EARLIER APPLICATION NUMBER: 60/093,209
; EARLIER FILING DATE: July 17, 1998
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 1
; LENGTH: 1858
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (10)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1138)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1175)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1201)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1342)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1344)
; US-09-347-819-1

Query Match 68.0%; Score 17; DB 3; Length 1858;
Best Local Similarity 80.0%; Pred. No. 2.3e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAACTGGATCTCAGGCTGAGA 25
||||| ||||| ||||| ||||| |||||
Db 1557 AAAAACTGGATCTCAGGCTGAGA 1581

RESULT 37
US-10-104-047-859
; Sequence 859, Application US/10104047
; Patent No. 6943241
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. 6943241el full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 859
; LENGTH: 2880
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-104-047-859


```

Db      18450 AAGAAACTAGAAATCGAGGCTGAGA 18474

RESULT 44
US-09-949-002-817
; Sequence 817, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL000790
; CURRENT APPLICATION NUMBER: US/09/949,002
; CURRENT FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/231,401
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 10823
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 817
; LENGTH: 49677
; TYPE: DNA
; ORGANISM: Human
US-09-949-002-817

Query Match      68.0%; Score 17; DB 3; Length 49677;
Best Local Similarity 80.0%; Pred. No. 3.9e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy      1 AAAAAACTGGAATCTCAGGCTGAGA 25
      ||||| ||||| ||||| ||||| ||||| |||||
Db      37426 AAAAAATTGGAACTGAGGCTAGGA 37450

RESULT 45
US-09-949-016-12458/c
; Sequence 12458, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12458
; LENGTH: 49818
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12458

Query Match      68.0%; Score 17; DB 3; Length 49818;
Best Local Similarity 80.0%; Pred. No. 3.9e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy      1 AAAAAACTGGAATCTCAGGCTGAGA 25
      ||||| ||||| ||||| ||||| ||||| |||||
Db      36549 AAAAACTAAATCTACGCTGAGA 36525

RESULT 46
US-09-949-016-14011/c
; Sequence 14011, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:

```

```
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14011
; LENGTH: 49829
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-14011

Query Match      68.0%; Score 17; DB 3; Length 49829;
Best Local Similarity 80.0%; Pred. No. 3.9e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY      1 AAAAACTGGAATCTCAGGCTGAGA 25
DB      36549 AAAAACTGGAATCTCAGGCTGAGA 36525

RESULT 47
US-09-949-016-12333
; Sequence 12333, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12333
; LENGTH: 85675
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12333

Query Match      68.0%; Score 17; DB 3; Length 85675;
Best Local Similarity 80.0%; Pred. No. 4.2e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY      1 AAAAACTGGAATCTCAGGCTGAGA 25
DB      55721 AAAAACTGGAATCTCAGGCTGAGA 55745

RESULT 48
US-09-949-016-15956
; Sequence 15956, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
```

```
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15956
; LENGTH: 85675
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15956

Query Match      68.0%; Score 17; DB 3; Length 85675;
Best Local Similarity 80.0%; Pred. No. 4.2e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY      1 AAAAACTGGAATCTCAGGCTGAGA 25
DB      55721 AAAAACTGGAATCTCAGGCTGAGA 55745

RESULT 49
US-09-949-016-16541/C
; Sequence 16541, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16541
; LENGTH: 96327
; TYPE: DNA
; ORGANISM: Human
; NAME/KEY: misc.feature
; LOCATION: (1)...(96327)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16541

Query Match      68.0%; Score 17; DB 3; Length 96327;
Best Local Similarity 80.0%; Pred. No. 4.3e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY      1 AAAAACTGGAATCTCAGGCTGAGA 25
DB      15778 AAAAACTGGAATCTCAGGCTGAGA 15754

RESULT 50
US-09-949-016-13029
; Sequence 13029, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
```

; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13029
; LENGTH: 105189
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(105189)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13029

Query Match 68.0%; Score 17; DB 3; Length 105189;
Best Local Similarity 80.0%; Pred. No. 4.3e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 AAAAACTGGAATCTCAGGCTGAGA 25
|||||
Db 17473 AAAAACTGGAATCTCAGGCTGAGA 17497
|||||

Search completed: February 3, 2006, 16:32:49
Job time : 88.6667 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 3, 2006, 14:37:28 ; Search time 365.556 Seconds
(without alignments)
565.535 Million cell updates/sec

Title: US-10-719-900-30

Perfect score: 25

Sequence: 1 aaaaactggaatctcaggctgaga 25

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database : Published Applications NA Main:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	25	100.0	25	8	US-10-719-900-30
2	23.4	93.6	25	8	US-10-719-900-29
3	20.2	80.8	583	4	US-09-925-065A-401109
4	20	80.0	300	9	US-10-779-543-1061
5	20	80.0	558	9	US-10-779-543-3989
6	18.8	75.2	537	4	US-09-925-065A-265072
7	18.8	75.2	63693	7	US-10-741-601-5650
8	18.6	74.4	589	4	US-09-925-065A-517136
9	18.6	74.4	2066	6	US-10-094-749-471
10	18.4	73.6	633	4	US-09-925-065A-417056
11	18.2	72.8	623	4	US-09-925-065A-637551
12	18.2	72.8	633	4	US-09-925-065A-523476
13	18.2	72.8	633	4	US-09-925-065A-523477
14	18.2	72.8	823	5	US-10-027-632-33065
15	18.2	72.8	823	6	US-10-027-632-33066
16	18.2	72.8	823	6	US-10-027-632-33065
17	18.2	72.8	823	6	US-10-027-632-33066
18	17.8	71.2	525	4	US-09-925-065A-362629
19	17.8	71.2	525	4	US-09-925-065A-362630
20	17.8	71.2	525	4	US-09-925-065A-362631
21	17.8	71.2	575	5	US-10-027-632-202264
22	17.8	71.2	575	6	US-10-027-632-202264
23	17.8	71.2	876	7	US-10-425-114-13369

7	US-10-424-599-113563	884	71.2	17.8	71.2	Sequence 113563,
1911	US-10-369-493-24074	1911	71.2	17.8	71.2	Sequence 24074, A
6	US-10-741-600-53948	201	70.4	17.6	70.4	Sequence 53948, A
214	US-09-933-797-614	214	70.4	17.6	70.4	Sequence 614, App
383	US-10-242-535A-41892	383	70.4	17.6	70.4	Sequence 41892, A
383	US-10-085-789A-41892	383	70.4	17.6	70.4	Sequence 41892, A
514	US-09-925-065A-311398	514	70.4	17.6	70.4	Sequence 311398,
4	US-09-925-065A-150366	516	70.4	17.6	70.4	Sequence 150366,
6	US-10-029-386-2155	522	70.4	17.6	70.4	Sequence 2155, Ap
538	US-09-925-065A-501583	538	70.4	17.6	70.4	Sequence 501583,
549	US-09-925-065A-435459	549	70.4	17.6	70.4	Sequence 435459,
549	US-09-925-065A-435470	549	70.4	17.6	70.4	Sequence 435470,
549	US-09-925-065A-435471	549	70.4	17.6	70.4	Sequence 435471,
549	US-09-925-065A-435473	549	70.4	17.6	70.4	Sequence 435473,
10	US-11-060-756-455	600	70.4	17.6	70.4	Sequence 455, App
600	US-11-060-756-456	600	70.4	17.6	70.4	Sequence 456, App
600	US-11-060-756-4727	600	70.4	17.6	70.4	Sequence 4727, Ap
604	US-11-060-756-4728	604	70.4	17.6	70.4	Sequence 4728, Ap
4	US-09-925-065A-288955	604	70.4	17.6	70.4	Sequence 288955,
4	US-09-925-065A-265864	606	70.4	17.6	70.4	Sequence 265864,
4	US-09-925-065A-563312	631	70.4	17.6	70.4	Sequence 563312,
5	US-10-027-632-191799	656	70.4	17.6	70.4	Sequence 191799,
6	US-10-027-632-191799	656	70.4	17.6	70.4	Sequence 191799,
4	US-09-925-065A-721185	1469	70.4	17.6	70.4	Sequence 721185,
7	US-10-424-599-123483	1551	70.4	17.6	70.4	Sequence 123483,
5	US-10-044-090-354	5271	70.4	17.6	70.4	Sequence 354, App
5	US-10-084-817-340	5271	70.4	17.6	70.4	Sequence 340, App
3	US-09-749-589-3	24526	70.4	17.6	70.4	Sequence 3, Appli
5	US-10-684-521-3	24526	70.4	17.6	70.4	Sequence 3, Appli
8	US-10-166-221-3	26320	70.4	17.6	70.4	Sequence 3, Appli
10	US-11-097-143-9055	37996	70.4	17.6	70.4	Sequence 9055, Ap
8	US-10-741-600-17758	51917	70.4	17.6	70.4	Sequence 17758, A
7	US-09-997-722-43	94	70.4	17.6	70.4	Sequence 43, Appl
7	US-10-450-826-105	154902	70.4	17.6	70.4	Sequence 105, App
7	US-10-322-696-16	299598	70.4	17.6	70.4	Sequence 16, Appl
5	US-10-087-192-1864	350764	70.4	17.6	70.4	Sequence 1864, Ap
5	US-10-087-192-226	653122	70.4	17.6	70.4	Sequence 226, App
4	US-09-925-065A-474511	580	69.6	17.4	69.6	Sequence 474511,
4	US-09-925-065A-783432	591	69.6	17.4	69.6	Sequence 783432,
5	US-10-027-632-169456	739	69.6	17.4	69.6	Sequence 169456,
739	US-10-027-632-169456	739	69.6	17.4	69.6	Sequence 169456,
7	US-10-425-114-219	1059	69.6	17.4	69.6	Sequence 219, App
7	US-10-425-114-2686	2206	69.6	17.4	69.6	Sequence 2686, Ap
7	US-10-425-114-34756	2806	69.6	17.4	69.6	Sequence 34756, A
5	US-10-425-115-143582	2899	69.6	17.4	69.6	Sequence 143582,
5	US-10-087-192-103	35456	69.6	17.4	69.6	Sequence 103, App
25	US-11-036-317-159236	25	68.8	17.2	68.8	Sequence 159236,
10	US-11-036-317-217422	25	68.8	17.2	68.8	Sequence 217422,
4	US-09-925-065A-604141	395	68.8	17.2	68.8	Sequence 604141,
4	US-09-925-065A-167140	428	68.8	17.2	68.8	Sequence 167140,
3	US-09-732-627A-3090	460	68.8	17.2	68.8	Sequence 3090, Ap
4	US-09-925-065A-150367	516	68.8	17.2	68.8	Sequence 150367,
4	US-09-925-065A-113663	518	68.8	17.2	68.8	Sequence 113663,
4	US-09-925-065A-113664	518	68.8	17.2	68.8	Sequence 113664,
4	US-09-925-065A-113665	518	68.8	17.2	68.8	Sequence 113665,
4	US-09-925-065A-522586	519	68.8	17.2	68.8	Sequence 522586,
4	US-09-925-065A-522588	519	68.8	17.2	68.8	Sequence 522588,
4	US-09-925-065A-647331	521	68.8	17.2	68.8	Sequence 647331,
5	US-10-027-632-180346	533	68.8	17.2	68.8	Sequence 180346,
5	US-10-027-632-299920	533	68.8	17.2	68.8	Sequence 299920,
6	US-10-027-632-180346	533	68.8	17.2	68.8	Sequence 180346,
6	US-10-027-632-299920	533	68.8	17.2	68.8	Sequence 299920,
5	US-10-027-632-76404	536	68.8	17.2	68.8	Sequence 76404, A
6	US-10-027-632-76404	536	68.8	17.2	68.8	Sequence 76404, A
6	US-10-027-632-228656	546	68.8	17.2	68.8	Sequence 228656,
6	US-10-027-632-228656	546	68.8	17.2	68.8	Sequence 228656,
4	US-09-925-065A-358845	547	68.8	17.2	68.8	Sequence 358845,
4	US-09-925-065A-14018	559	68.8	17.2	68.8	Sequence 14018, A
4	US-09-925-065A-312445	560	68.8	17.2	68.8	Sequence 312445,
4	US-09-925-065A-221346	569	68.8	17.2	68.8	Sequence 221346,
5	US-10-027-632-286926	569	68.8	17.2	68.8	Sequence 286926,
6	US-10-027-632-286926	569	68.8	17.2	68.8	Sequence 286926,
4	US-09-925-065A-221345	572	68.8	17.2	68.8	Sequence 221345,

```
C 97 17.2 68.8 573 5 US-10-027-632-84863
C 98 17.2 68.8 573 5 US-10-027-632-84864
C 99 17.2 68.8 573 5 US-10-027-632-286054
C 100 17.2 68.8 573 5 US-10-027-632-286055
C 101 17.2 68.8 573 6 US-10-027-632-84863
C 102 17.2 68.8 573 6 US-10-027-632-84864
C 103 17.2 68.8 573 6 US-10-027-632-286054
C 104 17.2 68.8 573 6 US-10-027-632-286055
C 105 17.2 68.8 589 4 US-09-925-065A-756159
C 106 17.2 68.8 598 4 US-09-925-065A-883540
C 107 17.2 68.8 604 4 US-09-925-065A-426028
C 108 17.2 68.8 664 5 US-10-027-632-242180
C 109 17.2 68.8 664 6 US-10-027-632-242180
C 110 17.2 68.8 672 4 US-09-925-065A-670810
C 111 17.2 68.8 693 4 US-09-925-065A-670811
C 112 17.2 68.8 693 8 US-10-425-115-15337
C 113 17.2 68.8 877 9 US-10-779-543-4813
C 114 17.2 68.8 944 8 US-10-767-795-5251
C 115 17.2 68.8 1184 4 US-09-925-065A-699263
C 116 17.2 68.8 1448 8 US-10-767-795-5252
C 117 17.2 68.8 2372 3 US-09-814-353-21640
C 118 17.2 68.8 3888 7 US-10-437-963-21899
C 119 17.2 68.8 4407 9 US-10-480-988-45
C 120 17.2 68.8 5054 8 US-10-723-860-5201
C 121 17.2 68.8 33486 6 US-10-085-959-143
C 122 17.2 68.8 35236 5 US-10-087-192-370
C 123 17.2 68.8 47841 7 US-10-433-580-3
C 124 17.2 68.8 76180 7 US-10-322-281-492
C 125 17.2 68.8 80959 3 US-09-858-546-3
C 126 17.2 68.8 143239 8 US-10-723-860-546
C 127 17.2 68.8 143239 9 US-10-756-149-566
C 128 17 68.0 136 3 US-09-815-242-2598
C 129 17 68.0 136 7 US-10-282-122A-5140
C 130 17 68.0 193 3 US-09-864-761-23226
C 131 17 68.0 229 7 US-10-424-599-6273
C 132 17 68.0 300 9 US-10-779-543-2758
C 133 17 68.0 349 3 US-09-770-791-721
C 134 17 68.0 388 9 US-10-779-543-16188
C 135 17 68.0 396 4 US-09-925-065A-272052
C 136 17 68.0 399 3 US-09-815-242-1432
C 137 17 68.0 399 7 US-10-282-122A-3996
C 138 17 68.0 402 3 US-09-960-352-8750
C 139 17 68.0 402 7 US-10-437-963-4507
C 140 17 68.0 406 7 US-10-424-599-124554
C 141 17 68.0 456 3 US-09-864-761-6513
C 142 17 68.0 464 8 US-10-357-930-9911
C 143 17 68.0 477 7 US-10-430-201-1124
C 144 17 68.0 477 8 US-10-430-201-1125
C 145 17 68.0 482 8 US-10-425-115-162668
C 146 17 68.0 486 3 US-09-918-993-10324
C 147 17 68.0 489 3 US-09-815-242-4216
C 148 17 68.0 501 3 US-09-815-242-8952
C 149 17 68.0 501 7 US-10-282-122A-7624
C 150 17 68.0 501 8 US-10-857-625-233
```

ALIGNMENTS

```
RESULT 1
US-10-719-900-30
; Sequence 30, Application US/10719900
; Publication No. US2005026164A1
; GENERAL INFORMATION:
; APPLICANT: Xue Mei Zhou
; TITLE OF INVENTION: Methods of Genetic Analysis of Mouse
; FILE REFERENCE: 3528.1
; CURRENT APPLICATION NUMBER: US/10719,900
; PRIOR FILING DATE: 2003-11-20
; PRIOR APPLICATION NUMBER: 60/427,808
; PRIOR FILING DATE: 2002 11 20
; NUMBER OF SEQ ID NOS: 982914
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1

; SEQ ID NO 30
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-719-900-30
Query Match 100.0%; Score 25; DB 8; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.084;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAAACTGGAATCTCAGGCTGAGA 25
|||||
Db 1 AAAAACTGGAATCTCAGGCTGAGA 25
|||||

RESULT 2
US-10-719-900-29
; Sequence 29, Application US/10719900
; Publication No. US2005026164A1
; GENERAL INFORMATION:
; APPLICANT: Xue Mei Zhou
; TITLE OF INVENTION: Methods of Genetic Analysis of Mouse
; FILE REFERENCE: 3528.1
; CURRENT APPLICATION NUMBER: US/10719,900
; PRIOR FILING DATE: 2003-11-20
; PRIOR APPLICATION NUMBER: 60/427,808
; PRIOR FILING DATE: 2002 11 20
; NUMBER OF SEQ ID NOS: 982914
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 29
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-719-900-29
Query Match 93.6%; Score 23.4; DB 8; Length 25;
Best Local Similarity 96.0%; Pred. No. 0.49;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAAACTGGAATCTCAGGCTGAGA 25
|||||
Db 1 AAAAACTGGAATCTCAGGCTGAGA 25
|||||

RESULT 3
US-09-925-065A-401109/c
; Sequence 401109, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 401109
; LENGTH: 583
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-401109
```

```
Query Match      80.0%; Score 20.2; DB 4; Length 583;
Best Local Similarity 88.0%; Pred. No. 26;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAACTGGAATCTCAGGCTGAGA 25
    |||||
Db 185 AGAAAACTGGAATCTCAGGCTGAAA 161

RESULT 4
US-10-779-543-1061
; Sequence 1061, Application US/10779543
; Publication No. US20050227917A1
; GENERAL INFORMATION:
; APPLICANT: Williams et al
; TITLE OF INVENTION: GENE PRODUCTS DIFFERENTIALLY EXPRESSED
; FILE REFERENCE: 2300-21302
; CURRENT APPLICATION NUMBER: US/10/779,543
; CURRENT FILING DATE: 2004-02-12
; PRIOR APPLICATION NUMBER: 10/076,555
; PRIOR FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: 09/217,471
; PRIOR FILING DATE: 1998-12-21
; PRIOR APPLICATION NUMBER: 60/068,755
; PRIOR FILING DATE: 1997-12-23
; PRIOR APPLICATION NUMBER: 60/080,664
; PRIOR FILING DATE: 1998-04-03
; PRIOR APPLICATION NUMBER: 60/105,234
; PRIOR FILING DATE: 1998-10-21
; PRIOR APPLICATION NUMBER: 09/297,648
; PRIOR FILING DATE: 2000-04-10
; PRIOR APPLICATION NUMBER: PCT/US99/01619
; PRIOR FILING DATE: 1999-01-28
; PRIOR APPLICATION NUMBER: 60/072,910
; PRIOR FILING DATE: 1998-01-28
; PRIOR APPLICATION NUMBER: 60/075,954
; PRIOR FILING DATE: 1998-02-24
; PRIOR APPLICATION NUMBER: 60/080,114
; PRIOR FILING DATE: 1998-03-31
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 23767
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1061
; LENGTH: 300
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-779-543-1061

Query Match      80.0%; Score 20; DB 9; Length 300;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 AAACTGGAATCTCAGGCTGA 23
    |||||
Db 104 AAACCTGGAATCTCAGGCTGA 123

RESULT 5
US-10-779-543-3989
; Sequence 3989, Application US/10779543
; Publication No. US20050227917A1
; GENERAL INFORMATION:
; APPLICANT: Williams et al
; TITLE OF INVENTION: GENE PRODUCTS DIFFERENTIALLY EXPRESSED
; FILE REFERENCE: 2300-21302
; CURRENT APPLICATION NUMBER: US/10/779,543
; CURRENT FILING DATE: 2004-02-12
; PRIOR APPLICATION NUMBER: 10/076,555
; PRIOR FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: 09/217,471
; PRIOR FILING DATE: 1998-12-21
```

```
; PRIOR APPLICATION NUMBER: 60/068,755
; PRIOR FILING DATE: 1997-12-23
; PRIOR APPLICATION NUMBER: 60/080,664
; PRIOR FILING DATE: 1998-04-03
; PRIOR APPLICATION NUMBER: 60/105,234
; PRIOR FILING DATE: 1998-10-21
; PRIOR APPLICATION NUMBER: 09/297,648
; PRIOR FILING DATE: 2000-04-10
; PRIOR APPLICATION NUMBER: PCT/US99/01619
; PRIOR FILING DATE: 1999-01-28
; PRIOR APPLICATION NUMBER: 60/072,910
; PRIOR FILING DATE: 1998-01-28
; PRIOR APPLICATION NUMBER: 60/075,954
; PRIOR FILING DATE: 1998-02-24
; PRIOR APPLICATION NUMBER: 60/080,114
; PRIOR FILING DATE: 1998-03-31
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 23767
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3989
; LENGTH: 758
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 5, 12, 15, 21, 39, 366, 370, 450, 460, 464, 536, 565, 567,
; LOCATION: 569, 586, 610, 627, 650, 666, 669, 686, 692, 695, 696, 720,
; LOCATION: 739, 743, 755
; OTHER INFORMATION: n = A,T,C or G
US-10-779-543-3989

Query Match      80.0%; Score 20; DB 9; Length 758;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 AAACCTGGAATCTCAGGCTGA 23
    |||||
Db 148 AAACCTGGAATCTCAGGCTGA 167

RESULT 6
US-09-925-065A-265072
; Sequence 265072, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 265072
; LENGTH: 537
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-265072

Query Match      75.2%; Score 18.8; DB 4; Length 537;
Best Local Similarity 90.9%; Pred. No. 1.2e+02;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

QY 3 AAAAAGTGAATCTCAGGCTGAG 24
 |||||
 Db 442 ACAACTGGAATCTCATGCTGAG 463

RESULT 7
 US-10-741-601-5650/c
 ; Sequence 5650, Application US/10741601
 ; Publication No. US20040166519A1
 ; GENERAL INFORMATION:
 ; APPLICANT: CARGILL, Michele et al.
 ; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
 ; STENOSIS, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001500
 ; CURRENT APPLICATION NUMBER: US/10/741,601
 ; CURRENT FILING DATE: 2003-12-22
 ; NUMBER OF SEQ ID NOS: 26415
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 5650
 ; LENGTH: 63693
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-741-601-5650

Query Match 75.2%; Score 18.8; DB 7; Length 63693;
 Best Local Similarity 90.9%; Pred. No. 2.5e+02;
 Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 AAAAAGTGAATCTCAGGCTGAG 24
 |||||
 Db 44895 ACAACTGGAATCTCATGCTGAG 44874

RESULT 8
 US-09-925-065A-517136
 ; Sequence 517136, Application US/09925065A
 ; Publication No. US20050228172A9
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, David G.
 ; TITLE OF INVENTION: Identification and Mapping of Single
 ; Nucleotide Polymorphisms in the Human Genome
 ; FILE REFERENCE: 108827.135
 ; CURRENT APPLICATION NUMBER: US/09/925,065A
 ; CURRENT FILING DATE: 2001-08-08
 ; PRIOR APPLICATION NUMBER: US 60/243,096
 ; PRIOR FILING DATE: 2000-10-24
 ; PRIOR APPLICATION NUMBER: US 60/252,147
 ; PRIOR FILING DATE: 2000-11-20
 ; PRIOR APPLICATION NUMBER: US 60/250,092
 ; PRIOR FILING DATE: 2000-11-30
 ; PRIOR APPLICATION NUMBER: US 60/261,766
 ; PRIOR FILING DATE: 2001-01-16
 ; PRIOR APPLICATION NUMBER: US 60/289,846
 ; PRIOR FILING DATE: 2001-05-09
 ; NUMBER OF SEQ ID NOS: 957086
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 517136
 ; LENGTH: 589
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-925-065A-517136

Query Match 74.4%; Score 18.6; DB 4; Length 589;
 Best Local Similarity 84.0%; Pred. No. 1.5e+02;
 Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAGTGAATCTCAGGCTGAG 25
 |||||
 Db 407 AAAGAAGTGAATCTCAGGCTGAG 431

RESULT 9
 US-10-094-749-471

; Sequence 471, Application US/10094749
 ; Publication No. US20030219741A1
 ; GENERAL INFORMATION:
 ; APPLICANT: ISOGAI, TAKAO
 ; APPLICANT: SUGIYAMA, TOMOYASU
 ; APPLICANT: OTSUKI, TETSUJI
 ; APPLICANT: WAKAMATSU, AI
 ; APPLICANT: SATO, HIROYUKI
 ; APPLICANT: ISHII, SHIZUKO
 ; APPLICANT: YAMAMOTO, JUN-ICHI
 ; APPLICANT: ISONO, YUUKO
 ; APPLICANT: HIO, YURI
 ; APPLICANT: OTSUKA, KAORU
 ; APPLICANT: NAGAI, KEIICHI
 ; APPLICANT: IRIE, RYOTARO
 ; APPLICANT: TAMECHIKA, ICHIRO
 ; APPLICANT: SEKI, NAOHICO
 ; APPLICANT: YOSHIKAWA, TSUTOMU
 ; APPLICANT: OTSUKA, MOTYUKI
 ; APPLICANT: NAGAHARI, KENJI
 ; APPLICANT: MASUHO, YASUHIKO
 ; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
 ; FILE REFERENCE: 084335/0160
 ; CURRENT APPLICATION NUMBER: US/10/094,749
 ; CURRENT FILING DATE: 2002-03-12
 ; PRIOR APPLICATION NUMBER: 60/350,435
 ; PRIOR FILING DATE: 2002-01-24
 ; PRIOR APPLICATION NUMBER: JP 2001-328381
 ; PRIOR FILING DATE: 2001-09-14
 ; NUMBER OF SEQ ID NOS: 3381
 ; SOFTWARE: Patent In Ver. 2.1
 ; SEQ ID NO 471
 ; LENGTH: 2066
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-094-749-471

Query Match 74.4%; Score 18.6; DB 6; Length 2066;
 Best Local Similarity 84.0%; Pred. No. 1.8e+02;
 Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAGTGAATCTCAGGCTGAG 25
 |||||
 Db 820 AAAAAGTGAATCTCAGGCTGAG 844

RESULT 10
 US-09-925-065A-417056
 ; Sequence 417056, Application US/09925065A
 ; Publication No. US20050228172A9
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, David G.
 ; TITLE OF INVENTION: Identification and Mapping of Single
 ; Nucleotide Polymorphisms in the Human Genome
 ; FILE REFERENCE: 108827.135
 ; CURRENT APPLICATION NUMBER: US/09/925,065A
 ; CURRENT FILING DATE: 2001-08-08
 ; PRIOR APPLICATION NUMBER: US 60/243,096
 ; PRIOR FILING DATE: 2000-10-24
 ; PRIOR APPLICATION NUMBER: US 60/252,147
 ; PRIOR FILING DATE: 2000-11-20
 ; PRIOR APPLICATION NUMBER: US 60/250,092
 ; PRIOR FILING DATE: 2000-11-30
 ; PRIOR APPLICATION NUMBER: US 60/261,766
 ; PRIOR FILING DATE: 2001-01-16
 ; PRIOR APPLICATION NUMBER: US 60/289,846
 ; PRIOR FILING DATE: 2001-05-09
 ; NUMBER OF SEQ ID NOS: 957086
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 417056
 ; LENGTH: 633
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens

```
;
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(633)
; OTHER INFORMATION: n = A,T,C or G
; US-09-925-065A-417056

Query Match      73.6%; Score 18.4; DB 4; Length 633;
Best Local Similarity 95.0%; Pred. No. 1.9e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY  2 AAAAAGTGAATCTCAGGCT 21
    ||||| ||||| ||||| |||||
Db   499 AAAATCTGAATCTCAGGCT 518

RESULT 11
US-09-925-065A-637551/c
; Sequence 637551, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 637551
; LENGTH: 623
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-925-065A-637551

Query Match      72.8%; Score 18.2; DB 4; Length 623;
Best Local Similarity 87.0%; Pred. No. 2.4e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY  3 AAAAAGTGAATCTCAGGCTGAGA 25
    ||||| ||||| ||||| |||||
Db   615 AAAAAGTGAATCTCAGGCTGAGA 593

RESULT 12
US-09-925-065A-523476/c
; Sequence 523476, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 523476
; LENGTH: 633
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-925-065A-523476

Query Match      72.8%; Score 18.2; DB 4; Length 633;
Best Local Similarity 87.0%; Pred. No. 2.4e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY  3 AAAAAGTGAATCTCAGGCTGAGA 25
    ||||| ||||| ||||| |||||
Db   49 AAAAAGTGAATCTCAGGCTGAGA 27

RESULT 14
US-10-027-632-33065
; Sequence 33065, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
```

; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 33065
; LENGTH: 823
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-33065

Query Match 72.8%; Score 18.2; DB 5; Length 823;
Best Local Similarity 87.0%; Pred. No. 2.5e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 AAAACTGGAATCTCAGGCTGACA 25
||||| |||||||||
Db 257 AAAACTACAATCTCAGGCTGACA 279

RESULT 15

US-10-027-632-33066
; Sequence 33066, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.

; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide

; FILE REFERENCE: 108827.129

; CURRENT APPLICATION NUMBER: US/10/027,632

; CURRENT FILING DATE: 2002-04-30

; PRIOR APPLICATION NUMBER: US 60/218,006

; PRIOR FILING DATE: 2000-07-12

; PRIOR APPLICATION NUMBER: US 60/198,676

; PRIOR FILING DATE: 2000-04-20

; PRIOR APPLICATION NUMBER: US 60/193,483

; PRIOR FILING DATE: 2000-03-29

; PRIOR APPLICATION NUMBER: US 60/185,218

; PRIOR FILING DATE: 2000-02-24

; PRIOR APPLICATION NUMBER: US 60/167,363

; PRIOR FILING DATE: 1999-11-23

; PRIOR APPLICATION NUMBER: US 60/156,358

; PRIOR FILING DATE: 1999-09-28

; PRIOR APPLICATION NUMBER: US 60/146,002

; PRIOR FILING DATE: 1999-08-09

; NUMBER OF SEQ ID NOS: 325720

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 33066

; LENGTH: 823

; TYPE: DNA

; ORGANISM: Human

US-10-027-632-33066

Query Match 72.8%; Score 18.2; DB 5; Length 823;
Best Local Similarity 87.0%; Pred. No. 2.5e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 AAAACTGGAATCTCAGGCTGACA 25
||||| |||||||||
Db 257 AAAACTACAATCTCAGGCTGACA 279

RESULT 16

US-10-027-632-33065
; Sequence 33065, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.

; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 33065
; LENGTH: 823
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-33065

Query Match 72.8%; Score 18.2; DB 6; Length 823;
Best Local Similarity 87.0%; Pred. No. 2.5e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 AAAACTGGAATCTCAGGCTGACA 25
||||| |||||||||
Db 257 AAAACTACAATCTCAGGCTGACA 279

RESULT 17

US-10-027-632-33066
; Sequence 33066, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.

; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide

; FILE REFERENCE: 108827.129

; CURRENT APPLICATION NUMBER: US/10/027,632

; CURRENT FILING DATE: 2002-04-30

; PRIOR APPLICATION NUMBER: US 60/218,006

; PRIOR FILING DATE: 2000-07-12

; PRIOR APPLICATION NUMBER: US 60/198,676

; PRIOR FILING DATE: 2000-04-20

; PRIOR APPLICATION NUMBER: US 60/193,483

; PRIOR FILING DATE: 2000-03-29

; PRIOR APPLICATION NUMBER: US 60/185,218

; PRIOR FILING DATE: 2000-02-24

; PRIOR APPLICATION NUMBER: US 60/167,363

; PRIOR FILING DATE: 1999-11-23

; PRIOR APPLICATION NUMBER: US 60/156,358

; PRIOR FILING DATE: 1999-09-28

; PRIOR APPLICATION NUMBER: US 60/146,002

; PRIOR FILING DATE: 1999-08-09

; NUMBER OF SEQ ID NOS: 325720

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 33066

; LENGTH: 823

; TYPE: DNA

; ORGANISM: Human

US-10-027-632-33066

Query Match 72.8%; Score 18.2; DB 6; Length 823;
Best Local Similarity 87.0%; Pred. No. 2.5e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 AAAAAGTGAATCTCAGGCTGAGA 25
||||| ||||| ||||| ||||| |||||
Db 257 AAAAAGTGAATCTCAGGCTGAGA 279

RESULT 18
US-09-925-065A-362629
; Sequence 362629, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 362629
; LENGTH: 525
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-362629

Query Match 71.2%; Score 17.8; DB 4; Length 525;
Best Local Similarity 90.5%; Pred. No. 3.6e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AAAAAGTGAATCTCAGGCTG 22
||||| ||||| ||||| ||||| |||||
Db 32 AAAAAGTGAATCTCAGGCTG 52

RESULT 19
US-09-925-065A-362630
; Sequence 362630, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 362630
; LENGTH: 525
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-362630

Query Match 71.2%; Score 17.8; DB 4; Length 525;

Best Local Similarity 90.5%; Pred. No. 3.6e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AAAAAGTGAATCTCAGGCTG 22
||||| ||||| ||||| ||||| |||||
Db 32 AAAAAGTGAATCTCAGGCTG 52

RESULT 20
US-09-925-065A-362631
; Sequence 362631, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 362631
; LENGTH: 525
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-362631

Query Match 71.2%; Score 17.8; DB 4; Length 525;
Best Local Similarity 90.5%; Pred. No. 3.6e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AAAAAGTGAATCTCAGGCTG 22
||||| ||||| ||||| ||||| |||||
Db 32 AAAAAGTGAATCTCAGGCTG 52

RESULT 21
US-10-027-632-202264/c
; Sequence 202264, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 202264
; LENGTH: 575
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-202264

Query Match 71.2%; Score 17.8; DB 5; Length 575;
Best Local Similarity 90.5%; Pred. No. 3.7e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AAAAACTGGAATCTCAGGCTG 22
||| ||||| ||||| |||||
Db 157 AAGAAGTGAATCTCAGGCTG 137

RESULT 22

US-10-027-632-202264/c
; Sequence 202264, Application US/10027632
; Publication No. US20030204075A9

GENERAL INFORMATION:

; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; POLYMORPHISMS in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1998-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 202264
; LENGTH: 575
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-202264

Query Match 71.2%; Score 17.8; DB 6; Length 575;
Best Local Similarity 90.5%; Pred. No. 3.7e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AAAAACTGGAATCTCAGGCTG 22
||| ||||| ||||| |||||
Db 157 AAGAAGTGAATCTCAGGCTG 137

RESULT 23

US-10-425-114-13369/c
; Sequence 13369, Application US/10425114
; Publication No. US2004003488A1

GENERAL INFORMATION:

; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; PLANT PRODUCTION
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 13369
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: JC-GMST02400057G10_FLI
US-10-425-114-13369

Query Match 71.2%; Score 17.8; DB 7; Length 876;
Best Local Similarity 90.5%; Pred. No. 3.9e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 AAACCTGGAATCTCAGGCTGAG 24
||||| ||||| ||||| |||||
Db 258 AAACCTGCACTCAGGCTGAG 238

RESULT 24

US-10-424-599-113563/c
; Sequence 113563, Application US/10424599
; Publication No. US20040031072A1

GENERAL INFORMATION:

; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; PLANT PRODUCTION
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 113563
; LENGTH: 884
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_73558C.1
US-10-424-599-113563

Query Match 71.2%; Score 17.8; DB 7; Length 884;
Best Local Similarity 90.5%; Pred. No. 3.9e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 AAACCTGGAATCTCAGGCTGAG 24
||||| ||||| ||||| |||||
Db 266 AAACCTGCAGTCTCAGGCTGAG 246

RESULT 25

US-10-369-493-24074
; Sequence 24074, Application US/10369493
; Publication No. US20030233675A1

GENERAL INFORMATION:

; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; PLANT PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 24074
; LENGTH: 1911
; TYPE: DNA
; ORGANISM: Xenorhabdus nematophilus
US-10-369-493-24074


```
Query Match          71.2%; Score 17.8; DB 6; Length 1911;
Best Local Similarity 90.5%; Pred. No. 4.4e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 AAAAAGTGAATCTCAGGCTGA 23
   |||||
Db 909 AAAAAGTGAATCTCAGGCTGA 929

RESULT 26
US-10-741-600-53948
; Sequence 53948, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 53948
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-741-600-53948

Query Match          70.4%; Score 17.6; DB 8; Length 201;
Best Local Similarity 83.3%; Pred. No. 3.9e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 AAAAAGTGAATCTCAGGCTGAG 24
   |||||
Db 102 AAAAAGTGAATCTCAGGCTGAG 125

RESULT 27
US-09-933-797-614/c
; Sequence 614, Application US/09933797
; Patent No. US20020155119A1
; GENERAL INFORMATION:
; APPLICANT: Robert A. Sikes et al.
; TITLE OF INVENTION: Isolation and Use of Fetal Urogenital
; Tissue
; FILE REFERENCE: 9901-007-999
; CURRENT APPLICATION NUMBER: US/09/933,797
; CURRENT FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: US/09/482,933
; PRIOR FILING DATE: 2000-01-14
; PRIOR APPLICATION NUMBER: PCT/US99/10746
; PRIOR FILING DATE: 1999-05/14
; PRIOR APPLICATION NUMBER: 60/085,383
; PRIOR FILING DATE: 1998-05-14
; NUMBER OF SEQ ID NOS: 811
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 614
; LENGTH: 214
; TYPE: DNA
; ORGANISM: Murine
US-09-933-797-614

Query Match          70.4%; Score 17.6; DB 3; Length 214;
Best Local Similarity 83.3%; Pred. No. 3.9e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 AAAAAGTGAATCTCAGGCTGAGA 25
   |||||
Db 58 AATCACTGGAGCTCAGGCTGAGA 35

RESULT 28
```

```
US-10-242-535A-41892/c
; Sequence 41892, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liew, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2005
; CURRENT APPLICATION NUMBER: US/10/242,535A
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 10/085,783
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 41892
; LENGTH: 383
; TYPE: DNA
; ORGANISM: Human
US-10-242-535A-41892

Query Match          70.4%; Score 17.6; DB 7; Length 383;
Best Local Similarity 83.3%; Pred. No. 4.3e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 AAAAAGTGAATCTCAGGCTGAG 24
   |||||
Db 141 AAAAAGTGAATCTCAGGCTGAG 118

RESULT 29
US-10-085-783A-41892/c
; Sequence 41892, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liew, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2002
; CURRENT APPLICATION NUMBER: US/10/085,783A
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 41892
; LENGTH: 383
; TYPE: DNA
; ORGANISM: Human
US-10-085-783A-41892

Query Match          70.4%; Score 17.6; DB 7; Length 383;
Best Local Similarity 83.3%; Pred. No. 4.3e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 AAAAAGTGAATCTCAGGCTGAG 24
   |||||
Db 141 AAAAAGTGAATCTCAGGCTGAG 118

RESULT 30
US-09-925-065A-311398
; Sequence 311398, Application US/09925065A
; Publication No. US20050228172A9
```

```
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 311398
; LENGTH: 514
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-311398
```

```
Query Match 70.4%; Score 17.6; DB 4; Length 514;
Best Local Similarity 83.3%; Pred. No. 4.5e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
```

```
QY 2 AAAAACTGGAATCTCAGGCTGAGA 25
||||| ||||| ||||| |||||
Db 247 AAAAGCTGGAGTCACAGGCTGAGA 270
```

RESULT 31

```
US-09-925-065A-150366/c
; Sequence 150366, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 150366
; LENGTH: 516
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-150366
```

```
Query Match 70.4%; Score 17.6; DB 4; Length 516;
Best Local Similarity 83.3%; Pred. No. 4.5e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
```

```
QY 1 AAAAACTGGAATCTCAGGCTGAG 24
||||| ||||| ||||| |||||
Db 215 AAAAACTGAAATGTAGGCTGTG 192
```

RESULT 32

```
US-10-029-386-2155/c
; Sequence 2155, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 2155
; LENGTH: 522
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR15.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.9
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 4.2
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.5
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.7
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 5.1
; OTHER INFORMATION: NT HIT: gi14756730, EVALUATE 1.00e-84
; OTHER INFORMATION: EST_HUMAN HIT: BE544267.1, EVALUATE 2.00e-84
US-10-029-386-2155
```

```
Query Match 70.4%; Score 17.6; DB 6; Length 522;
Best Local Similarity 83.3%; Pred. No. 4.5e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
```

```
QY 1 AAAAACTGGAATCTCAGGCTGAG 24
||||| ||||| ||||| |||||
Db 264 AAAAACTGAAATGTAGGCTGTG 241
```

RESULT 33

```
US-09-925-065A-501583
; Sequence 501583, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 501583
; LENGTH: 538
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-501583
```

```
Query Match 70.4%; Score 17.6; DB 4; Length 538;
Best Local Similarity 83.3%; Pred. No. 4.5e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
```

```
QY 1 AAAAACTGGAATCTCAGGCTGAG 24
||||| ||||| ||||| |||||
```

Db 272 AAAAATCTGGAATTCACAGGCTGAG 295

RESULT 34
US-09-925-065A-435469/c
; Sequence 435469, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 435469
; LENGTH: 549
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-435469

Query Match 70.4%; Score 17.6; DB 4; Length 549;
Best Local Similarity 83.3%; Pred. No. 4.5e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AAAAATCTGGAATTCACAGGCTGAGA 25
Db 417 AAAAATCTGGATCTCAGAAAGAGA 394

RESULT 35
US-09-925-065A-435470/c
; Sequence 435470, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 435470
; LENGTH: 549
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-435470

Query Match 70.4%; Score 17.6; DB 4; Length 549;
Best Local Similarity 83.3%; Pred. No. 4.5e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AAAAATCTGGAATTCACAGGCTGAGA 25
Db 417 AAAAATCTGGATCTCAGAAAGAGA 394

RESULT 36
US-09-925-065A-435471/c
; Sequence 435471, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 435471
; LENGTH: 549
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-435471

Query Match 70.4%; Score 17.6; DB 4; Length 549;
Best Local Similarity 83.3%; Pred. No. 4.5e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AAAAATCTGGAATTCACAGGCTGAGA 25
Db 417 AAAAATCTGGATCTCAGAAAGAGA 394

RESULT 37
US-09-925-065A-435473/c
; Sequence 435473, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 435473
; LENGTH: 549
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-435473

```
Query Match 70.4%; Score 17.6; DB 4; Length 549;
Best Local Similarity 83.3%; Pred. No. 4.5e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AAAAACTGGAATCTCAGGCTGAGA 25
||||| ||||| ||||| |||||
Db 417 AAAAACTGGGATCTCAGAAAGAGA 394

RESULT 38
US-11-060-756-455
; Sequence 455, Application US/11060756
; Publication No. US20050221354A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William Martin
; TITLE OF INVENTION: Nucleic Acid Arrays for Monitoring Expression Profiles of Drug
; FILE REFERENCE: AM101083 (031896-042000)
; CURRENT APPLICATION NUMBER: US/11/060,756
; CURRENT FILING DATE: 2005-02-18
; NUMBER OF SEQ ID NOS: 303284
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 455
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-060-756-455

Query Match 70.4%; Score 17.6; DB 10; Length 600;
Best Local Similarity 83.3%; Pred. No. 4.6e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AAAAACTGGAATCTCAGGCTGAGA 25
||||| ||||| ||||| |||||
Db 211 AAAAACTAGAACTCTGAAGGTGAGA 234

RESULT 39
US-11-060-756-456
; Sequence 456, Application US/11060756
; Publication No. US20050221354A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William Martin
; TITLE OF INVENTION: Nucleic Acid Arrays for Monitoring Expression Profiles of Drug
; FILE REFERENCE: AM101083 (031896-042000)
; CURRENT APPLICATION NUMBER: US/11/060,756
; CURRENT FILING DATE: 2005-02-18
; NUMBER OF SEQ ID NOS: 303284
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 456
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-060-756-456

Query Match 70.4%; Score 17.6; DB 10; Length 600;
Best Local Similarity 83.3%; Pred. No. 4.6e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AAAAACTGGAATCTCAGGCTGAGA 25
||||| ||||| ||||| |||||
Db 211 AAAAACTAGAACTCTGAAGGTGAGA 234

RESULT 40
US-11-060-756-4727
; Sequence 4727, Application US/11060756
; Publication No. US20050221354A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
```

```
; APPLICANT: Mounts, William Martin
; TITLE OF INVENTION: Nucleic Acid Arrays for Monitoring Expression Profiles of Drug
; FILE REFERENCE: AM101083 (031896-042000)
; CURRENT APPLICATION NUMBER: US/11/060,756
; CURRENT FILING DATE: 2005-02-18
; NUMBER OF SEQ ID NOS: 303284
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4727
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-060-756-4727

Query Match 70.4%; Score 17.6; DB 10; Length 600;
Best Local Similarity 83.3%; Pred. No. 4.6e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AAAAACTGGAATCTCAGGCTGAGA 25
||||| ||||| ||||| |||||
Db 211 AAAAACTAGAACTCTGAAGGTGAGA 234

RESULT 41
US-11-060-756-4728
; Sequence 4728, Application US/11060756
; Publication No. US20050221354A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William Martin
; TITLE OF INVENTION: Nucleic Acid Arrays for Monitoring Expression Profiles of Drug
; FILE REFERENCE: AM101083 (031896-042000)
; CURRENT APPLICATION NUMBER: US/11/060,756
; CURRENT FILING DATE: 2005-02-18
; NUMBER OF SEQ ID NOS: 303284
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4728
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-060-756-4728

Query Match 70.4%; Score 17.6; DB 10; Length 600;
Best Local Similarity 83.3%; Pred. No. 4.6e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AAAAACTGGAATCTCAGGCTGAGA 25
||||| ||||| ||||| |||||
Db 211 AAAAACTAGAACTCTGAAGGTGAGA 234

RESULT 42
US-09-925-065A-288955
; Sequence 288955, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
```


;
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 191799
; LENGTH: 656
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-191799

Query Match 70.4%; Score 17.6; DB 6; Length 656;
Best Local Similarity 83.3%; Pred. No. 4.7e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAACTGGAATCTCAGGCTGAG 24
||| ||||| ||||| ||||| |||||
DB 123 AAGAACTGCAATCTCATGCTGAG 100

RESULT 47
US-09-925-065A-721165
; Sequence 721165, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 721165
; LENGTH: 1469
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-721165

Query Match 70.4%; Score 17.6; DB 4; Length 1469;
Best Local Similarity 83.3%; Pred. No. 5.3e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AAAAACTGGAATCTCAGGCTGAGA 25
||| ||||| ||||| ||||| |||||
DB 491 AAATAATGGATTTCTGAGGCTGAGA 514

RESULT 48
US-10-424-599-123483/c
; Sequence 123483, Application US/10424599
; Publication No. US20040031072A1

;
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 123483
; LENGTH: 1551
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_82512C.1
US-10-424-599-123483

Query Match 70.4%; Score 17.6; DB 7; Length 1551;
Best Local Similarity 83.3%; Pred. No. 5.3e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AAAAACTGGAATCTCAGGCTGAGA 25
||| ||||| ||||| ||||| |||||
DB 1325 AAAAACTGGAATTTGAGGCAAGA 1302

RESULT 49
US-10-044-090-354/c
; Sequence 354, Application US/10044090
; Publication No. US20020137081A1
; GENERAL INFORMATION:
; APPLICANT: Olga Bandman
; TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION
; FILE REFERENCE: PA-0028 US
; CURRENT APPLICATION NUMBER: US/10/044,090
; CURRENT FILING DATE: 2002-01-09
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: PERL Program
; SEQ ID NO 354
; LENGTH: 5271
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20020137081A1 221500.1
; NAME/KEY: unsure
; LOCATION: 1086, 1089, 1773-2110
; OTHER INFORMATION: a, t, c, g, or other
US-10-044-090-354

Query Match 70.4%; Score 17.6; DB 5; Length 5271;
Best Local Similarity 83.3%; Pred. No. 6.4e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAACTGGAATCTCAGGCTGAG 24
||| ||||| ||||| ||||| |||||
DB 3345 AAATACTGGGATTACAGGCTGAG 3322

RESULT 50
US-10-084-817-340/c
; Sequence 340, Application US/10084817
; Publication No. US20030119009A1
; GENERAL INFORMATION:
; APPLICANT: Susan Stuart
; APPLICANT: Jed G. Nuchtern
; APPLICANT: Sharon E. Plon
; APPLICANT: Jason M. Shohet
; TITLE OF INVENTION: GENES REGULATED BY MYCN ACTIVATION
; FILE REFERENCE: PA-0046 US
; CURRENT APPLICATION NUMBER: US/10/084,817

Mon Feb 6 12:23:17 2006

;
; CURRENT FILING DATE: 2002-02-25
; PRIOR APPLICATION NUMBER: 60/270,784
; PRIOR FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 365
; SOFTWARE: PERL Program
; SEQ ID NO 340
; LENGTH: 5271
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030119009A1 221500.1
; NAME/KEY: unsure
; LOCATION: 1086, 1089, 1773-2110
; OTHER INFORMATION: a, t, c, g, or other
; US-10-084-817-340

Query Match 70.4%; Score 17.6; DB 5; Length 5271;
Best Local Similarity 83.3%; Pred. No. 6.4e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 AAAAACTGGATCTCAGGCTGAG 24
||| ||||| ||| ||||| |||
Db 3345 AAAATACGGATTACAGGCTGAG 3322

Search completed: February 3, 2006, 15:44:05
Job time : 368.556 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 3, 2006, 11:35:31 ; Search time 343.111 Seconds
(without alignments)
61.064 Million cell updates/sec

Title: US-10-719-900-30
Perfect score: 25
Sequence: 1 aaaaacggaatcagctgaga 25

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 6068529 seqs, 419036697 residues

Total number of hits satisfying chosen parameters: 12137058

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 150 summaries

Database : Published Applications NA New.*

- 1: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
- 2: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
- 5: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
- 7: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
- 9: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
- 10: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
- 11: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	18.8	75.2	63693	7 US-10-995-561-13269	Sequence 13269, A
2	18.6	74.4	682	7 US-10-750-185-58129	Sequence 58129, A
3	18.6	74.4	682	7 US-10-750-623-58129	Sequence 58129, A
4	18.6	74.4	1231	7 US-10-750-185-46140	Sequence 46140, A
5	18.6	74.4	1231	7 US-10-750-623-46140	Sequence 46140, A
6	18.4	73.6	187745	8 US-11-121-086-83	Sequence 83, Appl
7	17.6	70.4	201	7 US-10-995-561-58459	Sequence 58459, A
8	17.6	70.4	835	7 US-10-750-185-33984	Sequence 33984, A
9	17.6	70.4	835	7 US-10-750-623-33984	Sequence 33984, A
10	17.6	70.4	51917	7 US-10-995-561-13338	Sequence 63, Appl
11	17.4	69.6	169725	8 US-11-121-086-63	Sequence 1656, Ap
12	17.2	68.8	479	7 US-10-750-185-1656	Sequence 1656, Ap
13	17.2	68.8	479	7 US-10-750-623-1656	Sequence 1656, Ap
14	17.2	68.8	750	7 US-10-750-185-37889	Sequence 37889, A
15	17.2	68.8	750	7 US-10-750-623-37889	Sequence 37889, A
16	17.2	68.8	1768	7 US-10-750-185-32085	Sequence 32085, A
17	17.2	68.8	1768	7 US-10-750-623-32085	Sequence 32085, A
18	17	68.0	463	8 US-11-000-688-899	Sequence 899, App
19	17	68.0	475	8 US-11-000-688-899	Sequence 899, App
20	17	68.0	600	7 US-10-750-185-1913	Sequence 1913, Ap
21	17	68.0	600	7 US-10-750-623-1913	Sequence 1913, Ap
22	17	68.0	1094	8 US-11-136-527-285	Sequence 285, App

23	17	68.0	1094	8	US-11-136-527-4381	Sequence 4381, Ap
24	17	68.0	1412	7	US-10-750-185-40453	Sequence 40453, A
25	17	68.0	1412	7	US-10-750-623-40453	Sequence 40453, A
26	17	68.0	2084	7	US-10-485-517-61	Sequence 61, Appl
27	17	68.0	2084	7	US-10-485-517-76	Sequence 76, Appl
28	17	68.0	3272	7	US-10-750-185-31367	Sequence 31367, A
29	17	68.0	3272	7	US-10-750-623-31367	Sequence 31367, A
30	17	68.0	5796	7	US-10-821-234-62	Sequence 62, Appl
31	16.8	67.2	21	7	US-10-310-914A-626925	Sequence 626925, A
32	16.8	67.2	23	7	US-10-310-914A-626926	Sequence 626926, A
33	16.8	67.2	25	8	US-11-121-849-207753	Sequence 207753, A
34	16.8	67.2	25	8	US-11-136-527-207659	Sequence 207659, A
35	16.8	67.2	860	7	US-10-750-185-27917	Sequence 27917, A
36	16.8	67.2	2226	7	US-10-750-623-27917	Sequence 27917, A
37	16.8	67.2	2226	7	US-10-509-773-11	Sequence 11, Appl
38	16.6	66.4	600	7	US-10-750-185-2152	Sequence 2152, Ap
39	16.6	66.4	600	7	US-10-750-185-3019	Sequence 3019, Ap
40	16.6	66.4	600	7	US-10-750-623-2152	Sequence 2152, Ap
41	16.6	66.4	600	7	US-10-750-623-3019	Sequence 3019, Ap
42	16.6	66.4	1130	7	US-10-750-185-38194	Sequence 38194, A
43	16.6	66.4	1130	7	US-10-750-623-38194	Sequence 38194, A
44	16.6	66.4	2024	7	US-10-131-826A-359	Sequence 359, App
45	16.6	66.4	2078	7	US-10-750-185-37312	Sequence 37312, A
46	16.6	66.4	2078	7	US-10-750-623-37312	Sequence 37312, A
47	16.6	66.4	3969	7	US-10-750-185-53940	Sequence 53940, A
48	16.6	66.4	3969	7	US-10-750-623-53940	Sequence 53940, A
49	16.6	66.4	118996	8	US-11-121-086-84	Sequence 84, Appl
50	16.6	66.4	158410	8	US-11-121-086-46	Sequence 46, Appl
51	16.6	66.4	162013	8	US-11-150-888-30	Sequence 30, Appl
52	16.6	66.4	171936	7	US-10-933-023-24	Sequence 24, Appl
53	16.6	66.4	285300	7	US-10-857-780-6	Sequence 6, Appl
54	16.6	66.4	1082144	8	US-11-117-187-211	Sequence 211, App
55	16.6	66.4	1691140	8	US-11-091-018-1	Sequence 1, Appl
56	16.4	65.6	201	7	US-10-995-561-28189	Sequence 28189, A
57	16.4	65.6	600	7	US-10-750-185-1554	Sequence 1554, Ap
58	16.4	65.6	600	7	US-10-750-623-1554	Sequence 1554, Ap
59	16.4	65.6	5921	7	US-10-750-185-41152	Sequence 41152, A
60	16.4	65.6	5921	7	US-10-750-623-41152	Sequence 41152, A
61	16.4	65.6	222094	7	US-10-995-561-13244	Sequence 13244, A
62	16.2	64.8	25	8	US-11-121-849-320985	Sequence 320985, A
63	16.2	64.8	201	7	US-10-995-561-81886	Sequence 81886, A
64	16.2	64.8	201	7	US-10-995-561-81860	Sequence 81860, A
65	16.2	64.8	201	7	US-10-995-561-81961	Sequence 81961, A
66	16.2	64.8	679	7	US-10-750-185-53846	Sequence 53846, A
67	16.2	64.8	679	7	US-10-750-623-53846	Sequence 53846, A
68	16.2	64.8	922	7	US-10-750-185-36630	Sequence 36630, A
69	16.2	64.8	922	7	US-10-750-623-36630	Sequence 36630, A
70	16.2	64.8	1227	7	US-10-750-185-33375	Sequence 33375, A
71	16.2	64.8	1227	7	US-10-750-623-33375	Sequence 33375, A
72	16.2	64.8	1333	7	US-10-750-185-36500	Sequence 36500, A
73	16.2	64.8	1333	7	US-10-750-623-36500	Sequence 36500, A
74	16.2	64.8	1571	6	US-10-063-703-137	Sequence 137, App
75	16.2	64.8	1571	7	US-10-131-826A-97	Sequence 97, Appl
76	16.2	64.8	1571	7	US-10-131-826A-545	Sequence 545, App
77	16.2	64.8	1571	8	US-11-102-240-137	Sequence 137, App
78	16.2	64.8	1954	7	US-10-750-185-43339	Sequence 43339, A
79	16.2	64.8	1954	7	US-10-750-623-43339	Sequence 43339, A
80	16.2	64.8	6623	7	US-10-955-0548-149	Sequence 149, App
81	16.2	64.8	6623	8	US-11-121-086-67	Sequence 67, Appl
82	16.2	64.8	185393	8	US-11-121-086-101	Sequence 101, App
83	16.2	64.8	185393	8	US-11-121-086-78	Sequence 78, Appl
84	16.2	64.8	195235	7	US-10-995-561-13495	Sequence 13495, A
85	16.2	64.8	1691140	8	US-11-091-018-1	Sequence 1, Appl
86	16	64.0	25	8	US-11-121-849-282882	Sequence 282882, A
87	16	64.0	25	8	US-11-121-849-282970	Sequence 282970, A
88	16	64.0	201	8	US-11-124-368A-18177	Sequence 18177, A
89	16	64.0	201	8	US-11-124-368A-18482	Sequence 18482, A
90	16	64.0	201	8	US-11-124-368A-18527	Sequence 18527, A
91	16	64.0	433	8	US-11-000-688-322	Sequence 322, App
92	16	64.0	508	8	US-11-108-172-160	Sequence 160, App
93	16	64.0	768	7	US-10-750-185-43037	Sequence 43037, A
94	16	64.0	768	7	US-10-750-623-43037	Sequence 43037, A
95	16	64.0	992	7	US-10-750-185-53909	Sequence 53909, A

```
c 96 16 64.0 992 7 US-10-750-623-53909 Sequence 53909, A
c 97 16 64.0 1408 7 US-10-750-185-44804 Sequence 44804, A
c 98 16 64.0 1408 7 US-10-750-623-44804 Sequence 44804, A
c 99 16 64.0 1438 7 US-10-750-185-39949 Sequence 39949, A
c 100 16 64.0 1438 7 US-10-750-623-39949 Sequence 39949, A
c 101 16 64.0 1550 8 US-11-000-688-323 Sequence 323, App
c 102 16 64.0 1601 7 US-10-750-185-27423 Sequence 27423, A
c 103 16 64.0 1601 7 US-10-750-623-27423 Sequence 27423, A
c 104 16 64.0 1679 7 US-10-750-185-28409 Sequence 28409, A
c 105 16 64.0 1679 7 US-10-750-623-28409 Sequence 28409, A
c 106 16 64.0 1695 7 US-10-750-185-55993 Sequence 55993, A
c 107 16 64.0 1695 7 US-10-750-623-55993 Sequence 55993, A
c 108 16 64.0 1697 7 US-10-750-185-42862 Sequence 42862, A
c 109 16 64.0 1697 7 US-10-750-623-42862 Sequence 42862, A
c 110 16 64.0 1828 7 US-10-750-185-55873 Sequence 55873, A
c 111 16 64.0 1828 7 US-10-750-623-55873 Sequence 55873, A
c 112 16 64.0 1851 7 US-10-750-185-62541 Sequence 62541, A
c 113 16 64.0 1851 7 US-10-750-623-62541 Sequence 62541, A
c 114 16 64.0 1993 7 US-10-750-185-54772 Sequence 54772, A
c 115 16 64.0 1993 7 US-10-750-623-54772 Sequence 54772, A
c 116 16 64.0 2001 8 US-11-043-752-3895 Sequence 3895, Ap
c 117 16 64.0 2001 8 US-11-043-752-3895 Sequence 3895, Ap
c 118 16 64.0 2061 8 US-11-128-061-3438 Sequence 3438, Ap
c 119 16 64.0 2061 8 US-11-128-049-3438 Sequence 3438, Ap
c 120 16 64.0 2269 8 US-11-000-688-1374 Sequence 1374, Ap
c 121 16 64.0 2305 7 US-10-750-185-30133 Sequence 30133, A
c 122 16 64.0 2305 7 US-10-750-623-30133 Sequence 30133, A
c 123 16 64.0 2489 7 US-10-750-185-53608 Sequence 53608, A
c 124 16 64.0 2489 7 US-10-750-623-53608 Sequence 53608, A
c 125 16 64.0 2562 7 US-10-750-185-54101 Sequence 54101, A
c 126 16 64.0 2562 7 US-10-750-623-54101 Sequence 54101, A
c 127 16 64.0 2576 7 US-10-750-185-48750 Sequence 48750, A
c 128 16 64.0 2576 7 US-10-750-623-48750 Sequence 48750, A
c 129 16 64.0 2608 7 US-10-750-185-57741 Sequence 57741, A
c 130 16 64.0 2608 7 US-10-750-623-57741 Sequence 57741, A
c 131 16 64.0 2838 7 US-10-750-185-46440 Sequence 46440, A
c 132 16 64.0 2838 7 US-10-750-623-46440 Sequence 46440, A
c 133 16 64.0 2846 7 US-10-750-185-38714 Sequence 38714, A
c 134 16 64.0 2846 7 US-10-750-623-38714 Sequence 38714, A
c 135 16 64.0 3334 7 US-10-750-185-54572 Sequence 54572, A
c 136 16 64.0 3334 7 US-10-750-623-54572 Sequence 54572, A
c 137 16 64.0 3526 7 US-10-750-185-25056 Sequence 25056, A
c 138 16 64.0 3526 7 US-10-750-623-25056 Sequence 25056, A
c 139 16 64.0 3769 7 US-10-750-185-25076 Sequence 25076, A
c 140 16 64.0 3769 7 US-10-750-623-25076 Sequence 25076, A
c 141 16 64.0 3769 7 US-10-750-185-38543 Sequence 38543, A
c 142 16 64.0 3769 7 US-10-750-623-38543 Sequence 38543, A
c 143 16 64.0 3927 8 US-11-136-527-3008 Sequence 3008, Ap
c 144 16 64.0 5039 7 US-10-750-185-31557 Sequence 31557, A
c 145 16 64.0 5039 7 US-10-750-623-31557 Sequence 31557, A
c 146 16 64.0 18097 7 US-10-995-561-13486 Sequence 13486, A
c 147 16 64.0 18097 7 US-10-995-561-13486 Sequence 13486, A
c 148 16 64.0 20774 7 US-10-995-561-13291 Sequence 13291, A
c 149 16 64.0 23006 7 US-10-995-561-13465 Sequence 13465, A
c 150 16 64.0 96109 8 US-11-124-368A-2895 Sequence 2895, Ap
```

ALIGNMENTS

```
RESULT 1
US-10-995-561-13269/c
; Sequence 13269, Application US/10995561
; Publication No. US2005027054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702

Query Match 75.2%; Score 18.8; DB 7; Length 63693;
Best Local Similarity 90.9%; Pred. No. 44;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 AAAAAGTGAATCTCAGGCTGAG 24
| | | | | | | | | | | | | | | | | | | | | |
Db 44895 ACAACTGGAATCTCATGCTGAG 44874

RESULT 2
US-10-750-185-58129
; Sequence 58129, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 58129
; LENGTH: 682
; TYPE: DNA
; ORGANISM: Bovine 19866881916918
US-10-750-185-58129

Query Match 74.4%; Score 18.6; DB 7; Length 682;
Best Local Similarity 84.0%; Pred. No. 23;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAGTGAATCTCAGGCTGAGA 25
| | | | | | | | | | | | | | | | | | | | | |
Db 110 AAAAAGTGAATCTCTAGCTGAGA 134

RESULT 3
US-10-750-623-58129
; Sequence 58129, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 58129
; LENGTH: 682
```

```
; TYPE: DNA
; ORGANISM: Bovine 19866881916918
US-10-750-623-58129

Query Match          74.4%; Score 18.6; DB 7; Length 682;
Best Local Similarity 84.0%; Pred. No. 23;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAACTGGAATCTCAGGCTGAGA 25
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 110 AAAAAAGTGAATCTCTAGCTGAGA 134

RESULT 4
US-10-750-185-46140/c
; Sequence 46140, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 46140
; LENGTH: 1231
; TYPE: DNA
; ORGANISM: Bovine 19866881312657
US-10-750-185-46140

Query Match          74.4%; Score 18.6; DB 7; Length 1231;
Best Local Similarity 84.0%; Pred. No. 25;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAACTGGAATCTCAGGCTGAGA 25
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 651 AAGTAAGTGAATCTCAGGCTGAGA 627

RESULT 5
US-10-750-623-46140/c
; Sequence 46140, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 46140
; LENGTH: 1231
; TYPE: DNA
; ORGANISM: Bovine 19866881312657
US-10-750-623-46140

; TYPE: DNA
; ORGANISM: Bovine 19866881916918
US-10-750-623-58129

Query Match          74.4%; Score 18.6; DB 7; Length 1231;
Best Local Similarity 84.0%; Pred. No. 25;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAACTGGAATCTCAGGCTGAGA 25
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 651 AAGTAAGTGAATCTCAGGCTGAGA 627

RESULT 6
US-11-121-086-83
; Sequence 83, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: NIELSEN, KIRSTEN V.
; APPLICANT: POULSEN, TIM S.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138.6000-00000
; CURRENT APPLICATION NUMBER: US/11/121,086
; CURRENT FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: Patentin version 3.3
; SEQ ID NO 83
; LENGTH: 187745
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-83

Query Match          73.6%; Score 18.4; DB 8; Length 187745;
Best Local Similarity 95.0%; Pred. No. 84;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 AAAAACTGGAATCTCAGGCTG 22
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 15077 AAAAACTGGAATCTCAGGCTG 15096

RESULT 7
US-10-995-561-58459
; Sequence 58459, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 58459
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-58459

Query Match          70.4%; Score 17.6; DB 7; Length 201;
Best Local Similarity 83.3%; Pred. No. 52;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAACTGGAATCTCAGGCTGAG 24
   ||||| ||||| ||||| ||||| ||||| |||||
DB 102 AAAAACTGGAATCTCAGGCTGAG 125

RESULT 8
US-10-750-185-33984/c
; Sequence 33984, Application US/10750185
; Publication No. US20050260603A1
```

;
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; PRIOR FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 33984
; LENGTH: 835
; TYPE: DNA
; ORGANISM: Bovine 19866881966547
US-10-750-185-33984

Query Match 70.4%; Score 17.6; DB 7; Length 835;
Best Local Similarity 83.3%; Pred. No. 69;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AAAAACTGGAATCTCAGGCTGAGA 25
||||| ||||| ||||| ||||| |||||
Db 121 AAAAAATGGAATCTCCCGATGAGA 98

RESULT 9
US-10-750-623-33984/c
; Sequence 33984, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; PRIOR FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 33984
; LENGTH: 835
; TYPE: DNA
; ORGANISM: Bovine 19866881966547
US-10-750-623-33984

Query Match 70.4%; Score 17.6; DB 7; Length 835;
Best Local Similarity 83.3%; Pred. No. 69;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AAAAACTGGAATCTCAGGCTGAGA 25
||||| ||||| ||||| ||||| |||||
Db 121 AAAAAATGGAATCTCCCGATGAGA 98

RESULT 10
US-10-995-561-13338
; Sequence 13338, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH

;
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13338
; LENGTH: 51917
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)....(51917)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1.
US-10-995-561-13338

Query Match 70.4%; Score 17.6; DB 7; Length 51917;
Best Local Similarity 83.3%; Pred. No. 1.5e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAACTGGAATCTCAGGCTGAG 24
||||| ||||| ||||| ||||| |||||
Db 27042 AAAATACTGGAATTACAGGCTGAG 27065

RESULT 11
US-11-121-086-63/c
; Sequence 63, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; APPLICANT: NIELSEN, KIRSTEN V.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138.6000-00000
; CURRENT APPLICATION NUMBER: US/11/121,086
; CURRENT FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 63
; LENGTH: 169725
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-63

Query Match 69.6%; Score 17.4; DB 8; Length 169725;
Best Local Similarity 94.7%; Pred. No. 2.4e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 ACTGGAATCTCAGGCTGAG 24
||||| ||||| ||||| ||||| |||||
Db 110537 ACTGTAATCTCAGGCTGAG 110519

RESULT 12
US-10-750-185-1656/c
; Sequence 1656, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482

; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 1656
; LENGTH: 479
; TYPE: DNA
; ORGANISM: Bovine MMBT20262
US-10-750-185-1656

Query Match 68.8%; Score 17.2; DB 7; Length 479;
Best Local Similarity 86.4%; Pred. No. 94;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAACTGGAATCTCAGGCTGA 23
||||| | | | | | | | | | | | | | | | | |
DB 177 AAAACAGCACTCTCAGGCTGA 156

RESULT 13
US-10-750-623-1656/c
; Sequence 1656, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: FANTIN, Dennis
; APPLICANT: BATES, Stephen
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 1656
; LENGTH: 479
; TYPE: DNA
; ORGANISM: Bovine MMBT20262
US-10-750-623-1656

Query Match 68.8%; Score 17.2; DB 7; Length 479;
Best Local Similarity 86.4%; Pred. No. 94;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAACTGGAATCTCAGGCTGA 23
||||| | | | | | | | | | | | | | | | | |
DB 177 AAAACAGCACTCTCAGGCTGA 156

RESULT 14
US-10-750-185-37889/c
; Sequence 37889, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 1656
; LENGTH: 479
; TYPE: DNA
; ORGANISM: Bovine MMBT20262
US-10-750-623-1656

; SEQ ID NO 37889
; LENGTH: 750
; TYPE: DNA
; ORGANISM: Bovine 19866881213939
US-10-750-185-37889

Query Match 68.8%; Score 17.2; DB 7; Length 750;
Best Local Similarity 86.4%; Pred. No. 1e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAACTGGAATCTCAGGCTGA 23
||||| | | | | | | | | | | | | | | | | |
DB 448 AAAACAGCACTCTCAGGCTGA 427

RESULT 15
US-10-750-623-37889/c
; Sequence 37889, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: FANTIN, Dennis
; APPLICANT: BATES, Stephen
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 37889
; LENGTH: 750
; TYPE: DNA
; ORGANISM: Bovine 19866881213939
US-10-750-623-37889

Query Match 68.8%; Score 17.2; DB 7; Length 750;
Best Local Similarity 86.4%; Pred. No. 1e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAACTGGAATCTCAGGCTGA 23
||||| | | | | | | | | | | | | | | | | |
DB 448 AAAACAGCACTCTCAGGCTGA 427

RESULT 16
US-10-750-185-32085/c
; Sequence 32085, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 32085
; LENGTH: 1768
; TYPE: DNA
; ORGANISM: Bovine 19866881213939
US-10-750-623-37889

```
; ORGANISM: Bovine 19866880282614
US-10-750-185-32085

Query Match      68.8%; Score 17.2; DB 7; Length 1768;
Best Local Similarity 86.4%; Pred. No. 1.2e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAACTGGAATCTCAGGCTGA 23
   |||||||
Db 1464 AATGACTGGACTCTCAGGCTGA 1443

RESULT 17
US-10-750-623-32085/c
; Sequence 32085, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 32085
; LENGTH: 1768
; TYPE: DNA
; ORGANISM: Bovine 19866880282614
US-10-750-623-32085

Query Match      68.8%; Score 17.2; DB 7; Length 1768;
Best Local Similarity 86.4%; Pred. No. 1.2e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAACTGGAATCTCAGGCTGA 23
   |||||||
Db 1464 AATGACTGGACTCTCAGGCTGA 1443

RESULT 18
US-11-000-688-898/c
; Sequence 898, Application US/11000688
; Publication No. US20050287544A1
; GENERAL INFORMATION:
; APPLICANT: BERTUCCI, Francois
; APPLICANT: HOULGATTE, Remi
; APPLICANT: BIRNBAUM, Daniel
; TITLE OF INVENTION: GENE EXPRESSION PROFILING OF COLON CANCER WITH DNA ARRAYS
; FILE REFERENCE: 1423-R-03
; CURRENT APPLICATION NUMBER: US/11/000,688
; CURRENT FILING DATE: 2004-12-01
; PRIOR APPLICATION NUMBER: US 60/525,987
; PRIOR FILING DATE: 2003-12-01
; NUMBER OF SEQ ID NOS: 1596
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 898
; LENGTH: 463
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial sequences:primer
; NAME/KEY: misc_feature
; LOCATION: (1)-(463)
; OTHER INFORMATION: 3' terminal sequence from clone
; OTHER INFORMATION: image:19866880282614 gene.

; ORGANISM: Bovine 19866880282614
US-10-750-185-32085

Query Match      68.8%; Score 17; DB 8; Length 463;
Best Local Similarity 80.0%; Pred. No. 1.2e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAACTGGAATCTCAGGCTGAGA 25
   |||||||
Db 439 AAGTCACTGGAGGCTCAGGCTGAGA 415

RESULT 19
US-11-000-688-899
; Sequence 899, Application US/11000688
; Publication No. US20050287544A1
; GENERAL INFORMATION:
; APPLICANT: BERTUCCI, Francois
; APPLICANT: HOULGATTE, Remi
; APPLICANT: BIRNBAUM, Daniel
; TITLE OF INVENTION: GENE EXPRESSION PROFILING OF COLON CANCER WITH DNA ARRAYS
; FILE REFERENCE: 1423-R-03
; CURRENT APPLICATION NUMBER: US/11/000,688
; CURRENT FILING DATE: 2004-12-01
; PRIOR APPLICATION NUMBER: US 60/525,987
; PRIOR FILING DATE: 2003-12-01
; NUMBER OF SEQ ID NOS: 1596
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 899
; LENGTH: 475
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial sequences:primer
; NAME/KEY: misc_feature
; LOCATION: (1)-(475)
; OTHER INFORMATION: 5' terminal sequence from clone
; OTHER INFORMATION: image:358683.tropomyosin 1 (alpha) (TPM1) gene.

; ORGANISM: Bovine 19866880282614
US-10-750-623-32085

Query Match      68.8%; Score 17.2; DB 7; Length 1768;
Best Local Similarity 86.4%; Pred. No. 1.2e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAACTGGAATCTCAGGCTGA 23
   |||||||
Db 1464 AATGACTGGACTCTCAGGCTGA 1443

RESULT 18
US-11-000-688-898/c
; Sequence 898, Application US/11000688
; Publication No. US20050287544A1
; GENERAL INFORMATION:
; APPLICANT: BERTUCCI, Francois
; APPLICANT: HOULGATTE, Remi
; APPLICANT: BIRNBAUM, Daniel
; TITLE OF INVENTION: GENE EXPRESSION PROFILING OF COLON CANCER WITH DNA ARRAYS
; FILE REFERENCE: 1423-R-03
; CURRENT APPLICATION NUMBER: US/11/000,688
; CURRENT FILING DATE: 2004-12-01
; PRIOR APPLICATION NUMBER: US 60/525,987
; PRIOR FILING DATE: 2003-12-01
; NUMBER OF SEQ ID NOS: 1596
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 898
; LENGTH: 463
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial sequences:primer
; NAME/KEY: misc_feature
; LOCATION: (1)-(463)
; OTHER INFORMATION: 3' terminal sequence from clone
; OTHER INFORMATION: image:19866880282614 gene.

; ORGANISM: Bovine 19866880282614
US-10-750-185-32085

Query Match      68.8%; Score 17; DB 8; Length 475;
Best Local Similarity 80.0%; Pred. No. 1.2e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAACTGGAATCTCAGGCTGAGA 25
   |||||||
Db 411 AAGTCACTGGAGGCTCAGGCTGAGA 435

RESULT 20
US-10-750-185-1913
; Sequence 1913, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
```

```

; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 1913
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Bovine MMBT18521
US-10-759-185-1913

```

```

Query Match      68.0%; Score 17; DB 7; Length 600;
Best Local Similarity 80.0%; Pred. No. 1.2e+02;
Matches 20; Conservative 0; Mismatches 5; Indels

QY 1 AAAAAA CTGGAATCTCAGCTGAGA 25
    |||
Db 189 AACATAC TCGAATCTCTGGCTGAGA 213
    |||

```

```

RESULT 21
US-10-750-623-1913
; Sequence 1913, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC..
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 1913
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Bovine MBMT18521
US-10-750-623-1913

```

Query Match	68.0%	Score 17	DB 7	Length 600
Best Local Similarity	80.0%	Pred. No. 1.2e+02		
Matches 20	Conservative 0	Mismatches 5	Indels 0	Gaps 0
Qy	1	AAAAAACTGGAATCTCAGGCTGAGA	25	
Dh	189	AACATACTGGAATCTCGGCTGGA	213	

```

RESULT 22
US-11-136-527-285
; Sequence 285, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 285
; LENGTH: 1094
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-285

```

Query Match 68.0%; Score 17; DB 8; Length 1094;

Best Local Similarity 80.0%; Pred. No. 1.4e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 AAAAACTGGAAATCTCAGGCTGAGA 25
||| ||| ||| ||| ||| ||| |||
Db 641 AAGTCACTGGAGGCTCAGGCTGAGA 665

```

RESULT 23
US-11-136-527-4381
; Sequence 4381, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Wyeth
; APPLICANT: William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AMI101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4381
; LENGTH: 1094
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-4381

```

```

Query Match      68.0%; Score 17; DB 8; Length 1094;
Best Local Similarity 80.0%; Pred.No. 1.4e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 1 AAAAAACTGGAATCTTCAGGCTGAGA 25
    ||| ||||| ||||| ||||| |||||
Db 641 AACTCACTGAGGCTCAGGCTGAGA 665

```

```

RESULT 24
US-10-750-185-40453/c
; Sequence 40453, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 40453
; LENGTH: 1412
; TYPE: DNA
; ORGANISM: Bovine 19866880538167
US-10-750-185-40453

```

```
Query Match      68.0%; Score 17; DB 7; Length 1412;
Best Local Similarity 80.0%; Pred. No. 1.4e+02;
Matches 20; Conservative 0; Mismatches 5; Indels

QY    1 AAAAAACTGGAATCTCAGGCTGAGA 25
      ||| ||| ||| ||| ||| ||| |||
nb    363 AAACACCTTGAACAATTAGGCTTGAGA 339
```

```
RESULT 25
US-10-750-623-40453/c
; Sequence 40453, Application US/10750623
; Publication NO. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 40453
; LENGTH: 1412
; TYPE: DNA
; ORGANISM: Bovine 19866880538167
US-10-750-623-40453

Query Match      68.0%; Score 17; DB 7; Length 1412;
Best Local Similarity 80.0%; Pred. No. 1.4e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAACTGGAATCTCAGGCTGAGA 25
Db 363 AAAGACCTAGAAACTTAGGCTGAGA 339

RESULT 26
US-10-485-517-61
; Sequence 61, Application US/10485517
; Publication NO. US20050256299A1
; GENERAL INFORMATION:
; APPLICANT: University of Sheffield
; APPLICANT: BIOSYNEXUS INCORPORATED
; APPLICANT: FOSTER, Simon
; APPLICANT: MOND, James
; TITLE OF INVENTION: Antigenic Polypeptides
; FILE REFERENCE: P100629W0
; CURRENT APPLICATION NUMBER: US/10/485,517
; CURRENT FILING DATE: 2004-02-02
; PRIOR APPLICATION NUMBER: GB 0118825.9
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: GB 0200349.9
; PRIOR FILING DATE: 2002-01-09
; NUMBER OF SEQ ID NOS: 424
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 61
; LENGTH: 2084
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-10-485-517-61

Query Match      68.0%; Score 17; DB 7; Length 2084;
Best Local Similarity 80.0%; Pred. No. 1.6e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAACTGGAATCTCAGGCTGAGA 25
Db 396 AACATCCGGAATCTTAATGCTGAGA 420

RESULT 27
US-10-485-517-76
; Sequence 76, Application US/10485517
; Publication NO. US20050256299A1
```

```
; GENERAL INFORMATION:
; APPLICANT: University of Sheffield
; APPLICANT: Biosynexus Incorporated
; APPLICANT: Foster, Simon
; APPLICANT: Mond, James
; TITLE OF INVENTION: Antigenic Polypeptides
; FILE REFERENCE: P100629W0
; CURRENT APPLICATION NUMBER: US/10/485,517
; CURRENT FILING DATE: 2004-02-02
; PRIOR APPLICATION NUMBER: GB 0118825.9
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: GB 0200349.9
; PRIOR FILING DATE: 2002-01-09
; NUMBER OF SEQ ID NOS: 424
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 76
; LENGTH: 2084
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-10-485-517-76

Query Match      68.0%; Score 17; DB 7; Length 2084;
Best Local Similarity 80.0%; Pred. No. 1.6e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAACTGGAATCTCAGGCTGAGA 25
Db 396 AACATCCGGAATCTTAATGCTGAGA 420

RESULT 28
US-10-750-185-31367
; Sequence 31367, Application US/10750185
; Publication NO. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 31367
; LENGTH: 3272
; TYPE: DNA
; ORGANISM: Bovine 19866881031011
US-10-750-185-31367

Query Match      68.0%; Score 17; DB 7; Length 3272;
Best Local Similarity 80.0%; Pred. No. 1.7e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAACTGGAATCTCAGGCTGAGA 25
Db 1804 AAAATAGGAATCTGAGGCTGAGA 1828

RESULT 29
US-10-750-623-31367
; Sequence 31367, Application US/10750623
; Publication NO. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
```



```
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 31367
; LENGTH: 3272
; TYPE: DNA
; ORGANISM: Bovine 19866881031011
US-10-750-623-31367

Query Match      68.0%; Score 17; DB 7; Length 3272;
Best Local Similarity 80.0%; Pred. No. 1.7e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAACTGGAATCTCAGGCTGAGA 25
    ||||| ||||| ||||| |||||
Db 1804 AAAAAAAGGAATCTCAGGCTCAGA 1828

RESULT 30
US-10-821-234-62/c
; Sequence 62, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pt_seq_genes Version 1.0
; SEQ ID NO 62
; LENGTH: 5796
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-821-234-62

Query Match      68.0%; Score 17; DB 7; Length 5796;
Best Local Similarity 80.0%; Pred. No. 1.9e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAACTGGAATCTCAGGCTGAGA 25
    ||||| ||||| ||||| |||||
Db 3003 AAGAAGTGGACCTCGAGTTGAGA 2979

RESULT 31
US-10-310-914A-626925
; Sequence 626925, Application US/10310914A
; Publication No. US20060003322A1
; GENERAL INFORMATION:
; APPLICANT: Bentwich, Isaac
; APPLICANT: Shiler, Kvuzat
; TITLE OF INVENTION: Bioinformatically detectable group of novel regulatory genes and
; TITLE OF INVENTION: uses thereof
; FILE REFERENCE: 06087.0200.CPUS01
; CURRENT APPLICATION NUMBER: US/10/310,914A
; CURRENT FILING DATE: 2002-12-06
; NUMBER OF SEQ ID NOS: 1388402
; SOFTWARE: PatentIn version 3.3
```

```
; SEQ ID NO 626925
; LENGTH: 21
; TYPE: RNA
; ORGANISM: Human
US-10-310-914A-626925

Query Match      67.2%; Score 16.8; DB 7; Length 21;
Best Local Similarity 70.0%; Pred. No. 78;
Matches 14; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 4 AAACGTGAATCTCAGGCTGA 23
    ||| :||| :||| :||| :|||
Db 2 AACCGGACUCUCAGGCUGA 21

RESULT 32
US-10-310-914A-626926
; Sequence 626926, Application US/10310914A
; Publication No. US20060003322A1
; GENERAL INFORMATION:
; APPLICANT: Bentwich, Isaac
; APPLICANT: Shiler, Kvuzat
; TITLE OF INVENTION: Bioinformatically detectable group of novel regulatory genes and
; TITLE OF INVENTION: uses thereof
; FILE REFERENCE: 06087.0200.CPUS01
; CURRENT APPLICATION NUMBER: US/10/310,914A
; CURRENT FILING DATE: 2002-12-06
; NUMBER OF SEQ ID NOS: 1388402
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 626926
; LENGTH: 23
; TYPE: RNA
; ORGANISM: Human
US-10-310-914A-626926

Query Match      67.2%; Score 16.8; DB 7; Length 23;
Best Local Similarity 70.0%; Pred. No. 80;
Matches 14; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 4 AAACGTGAATCTCAGGCTGA 23
    ||| :||| :||| :||| :|||
Db 2 AACCGGACUCUCAGGCUGA 21

RESULT 33
US-11-121-849-207753/c
; Sequence 207753, Application US/11121849
; Publication No. US20050272080A1
; GENERAL INFORMATION:
; APPLICANT: John Palma
; TITLE OF INVENTION: Methods of Genetic Analysis of Formalin Fixed Paraffin Embedded
; TITLE OF INVENTION: Microarrays
; FILE REFERENCE: 3684.1
; CURRENT APPLICATION NUMBER: US/11/121,849
; CURRENT FILING DATE: 2005-05-03
; PRIOR APPLICATION NUMBER: 60/567,949
; PRIOR FILING DATE: 2004-05-03
; NUMBER OF SEQ ID NOS: 673904
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 207753
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapien
US-11-121-849-207753

Query Match      67.2%; Score 16.8; DB 8; Length 25;
Best Local Similarity 90.0%; Pred. No. 81;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 AAACGTGAATCTCAGGCTGA 23
    ||| ||||| ||||| ||||| |||||
Db 21 AACCTGGACTCTCAGGCTGA 2
```

RESULT 34
US-11-136-527-207659
; Sequence 207659, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 207659
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Probe
US-11-136-527-207659

Query Match 67.2%; Score 16.8; DB 8; Length 25;
Best Local Similarity 90.0%; Pred. No. 81;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 ACTGGAATCTCAGGCTGAGA 25
||||| ||||| ||||| |||||
Db 3 ACTGGAGGCTCAGGCTGAGA 22

RESULT 35
US-10-750-185-27917/c
; Sequence 27917, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 27917
; LENGTH: 860
; TYPE: DNA
; ORGANISM: Bovine 19866881073914
US-10-750-185-27917

Query Match 67.2%; Score 16.8; DB 7; Length 860;
Best Local Similarity 90.0%; Pred. No. 1.6e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AAAAACTGGAATCTCAGGCT 21
||||| ||||| ||||| |||||
Db 603 AAAAATGGCATCTCAGGTT 584

RESULT 36
US-10-750-623-27917/c
; Sequence 27917, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: DENISE, Sue K.

; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 27917
; LENGTH: 860
; TYPE: DNA
; ORGANISM: Bovine 19866881073914
US-10-750-623-27917

Query Match 67.2%; Score 16.8; DB 7; Length 860;
Best Local Similarity 90.0%; Pred. No. 1.6e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AAAAACTGGAATCTCAGGCT 21
||||| ||||| ||||| |||||
Db 603 AAAAATGGCATCTCAGGTT 584

RESULT 37
US-10-509-773-11/c
; Sequence 11, Application US/10509773
; Publication No. US20050281743A1
; GENERAL INFORMATION:
; APPLICANT: Delaney, Allen
; TITLE OF INVENTION: Cancer Associated Protein Phosphatases and their
; FILE REFERENCE: SMAR-044
; CURRENT APPLICATION NUMBER: US/10/509,773
; CURRENT FILING DATE: 2004-09-28
; PRIOR APPLICATION NUMBER: CA03/00393
; PRIOR FILING DATE: 2003-03-19
; PRIOR APPLICATION NUMBER: 60/368,859
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 2226
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: FLJ20313 nucleotide sequence
US-10-509-773-11

Query Match 67.2%; Score 16.8; DB 7; Length 2226;
Best Local Similarity 90.0%; Pred. No. 2e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 AAACCTGGAATCTCAGGCTGA 23
||||| ||||| ||||| |||||
Db 2057 AACCTGGACTCTCAGGCTGA 2038

RESULT 38
US-10-750-185-2152/c
; Sequence 2152, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.

; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 2152
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Bovine MMBT21073
US-10-750-185-2152

Query Match 66.4%; Score 16.6; DB 7; Length 600;
Best Local Similarity 82.6%; Pred. No. 1.9e+02;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 AAAACTGGAATCTCAGGCTGAGA 25
||||| ||||| ||||| ||||| |||||
Db 202 AAAACTGCAATCTCATCTTGAGA 180

RESULT 39
US-10-750-185-3019
; Sequence 3019, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 3019
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Bovine MMBT21143
US-10-750-185-3019

Query Match 66.4%; Score 16.6; DB 7; Length 600;
Best Local Similarity 82.6%; Pred. No. 1.9e+02;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 AAAACTGGAATCTCAGGCTGAGA 25
||||| ||||| ||||| ||||| |||||
Db 91 AGAACAGGCATCTCAGGCTCAGA 113

RESULT 40
US-10-750-623-2152/c
; Sequence 2152, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom

; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 2152
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Bovine MMBT21073
US-10-750-623-2152

Query Match 66.4%; Score 16.6; DB 7; Length 600;
Best Local Similarity 82.6%; Pred. No. 1.9e+02;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 AAAACTGGAATCTCAGGCTGAGA 25
||||| ||||| ||||| ||||| |||||
Db 202 AAAACTGCAATCTCATCTTGAGA 180

RESULT 41
US-10-750-623-3019
; Sequence 3019, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 3019
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Bovine MMBT21143
US-10-750-623-3019

Query Match 66.4%; Score 16.6; DB 7; Length 600;
Best Local Similarity 82.6%; Pred. No. 1.9e+02;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 AAAACTGGAATCTCAGGCTGAGA 25
||||| ||||| ||||| ||||| |||||
Db 91 AGAACAGGCATCTCAGGCTCAGA 113

RESULT 42
US-10-750-185-38194
; Sequence 38194, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS

```

; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 38194
; LENGTH: 1130
; TYPE: DNA
; ORGANISM: Bovine 19866880591492
US-10-750-185-38194

```

```

Query Match 66.4%; Score 16.6; DB 7; Length 1130;
Best Local Similarity 82.6%; Pred. No. 2.1e+02;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

```

QY 3 AAAAAGGATCTCAGGCTGAGA 25
    ||||| ||||| ||||| |||||
Db 541 AGACAGGATCTCAGGCTCAGA 563

```

```

RESULT 43
US-10-750-623-38194
; Sequence 38194, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 38194
; LENGTH: 1130
; TYPE: DNA
; ORGANISM: Bovine 19866880591492
US-10-750-623-38194

```

```

Query Match 66.4%; Score 16.6; DB 7; Length 1130;
Best Local Similarity 82.6%; Pred. No. 2.1e+02;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

```

QY 3 AAAAAGGATCTCAGGCTGAGA 25
    ||||| ||||| ||||| |||||
Db 541 AGACAGGATCTCAGGCTCAGA 563

```

```

RESULT 44
US-10-131-826A-259
; Sequence 259, Application US/10131826A
; Publication No. US20050245730A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven

```

```

; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C128
; CURRENT APPLICATION NUMBER: US/10/131,826A
; CURRENT FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: 60/049911
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059117
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059122
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059184
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059588
; PRIOR FILING DATE: 1997-09-19
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 259
; LENGTH: 2024
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-131-826A-259

```

```

Query Match 66.4%; Score 16.6; DB 7; Length 2024;
Best Local Similarity 82.6%; Pred. No. 2.4e+02;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

```

QY 1 AAAAAGGATCTCAGGCTGA 23
    ||||| ||||| ||||| |||||
Db 1419 AACAACTGGAATTCATCTGA 1441

```

```

RESULT 45
US-10-750-185-37312/c
; Sequence 37312, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 37312
; LENGTH: 2078
; TYPE: DNA
; ORGANISM: Bovine 19866880966268
US-10-750-185-37312

```



```

US-11-121-086-46
; Sequence 46, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138.6000-00000
; CURRENT APPLICATION NUMBER: US/11/121,086
; PRIOR FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 46
; LENGTH: 158410
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-46

Query Match      66.4%; Score 16.6; DB 8; Length 158410;
Best Local Similarity 82.6%; Pred. No. 5.3e+02;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      3 AAAACTGGAATCTCAGGCTGAGA 25
Db      148716 AAATCTGGAATCCAGCGGAGA 148738

Search completed: February 3, 2006, 16:20:21
Job time : 347.111 secs

```

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 3, 2006, 10:53:23 ; Search time 863 Seconds
(without alignments)
1646.682 Million cell updates/sec

Title: US-10-719-900-35

Perfect score: 25

Sequence: 1 aaaaagagcgagactgtcttc 25

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database :

GenEmbl:*

1: gb_ba:*

2: gb_in:*

3: gb_env:*

4: gb_on:*

5: gb_ov:*

6: gb_pat:*

7: gb_ph:*

8: gb_pr:*

9: gb_ro:*

10: gb_sts:*

11: gb_sy:*

12: gb_un:*

13: gb_vi:*

14: gb_htg:*

15: gb_pl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	23.4	93.6	163878	9	AC134563 Mus muscu
2	23.4	93.6	193811	9	AC142098 Mus muscu
3	21.8	87.2	28365	14	AC166153 Mus muscu
4	21.8	87.2	98311	9	CR788287 Mouse DNA
5	21.8	87.2	110000	14	Continuation (2 of BX546458_1
6	21.8	87.2	110000	14	BX546462_0
7	21.8	87.2	110000	14	CR753898 Mus muscu
8	21.8	87.2	110000	14	Continuation (3 of BX284639_2
9	21.8	87.2	115602	9	BX284639 Mouse DNA
10	21.8	87.2	136564	14	CR974476 Mus muscu
11	21.8	87.2	143116	14	AC118310 Rattus no
12	21.8	87.2	143176	14	AC166152 Mus muscu
13	21.8	87.2	145110	9	CR354442 Mouse DNA
14	21.8	87.2	145110	9	CR354442 Mouse DNA
15	21.8	87.2	145376	14	CT010482 CT010482 Mus muscu
16	21.8	87.2	149369	9	AL607123 Mouse DNA
17	21.8	87.2	151693	14	CR955032 Homo sapi
18	21.8	87.2	169581	9	BX511235 Mouse DNA

19	21.8	87.2	174479	14	CR955031 Homo sapi
20	21.8	87.2	174479	14	CR955031 Homo sapi
21	21.8	87.2	179228	9	BX005149 Mouse DNA
22	21.8	87.2	179228	9	BX005149 Mouse DNA
23	21.8	87.2	180947	9	AL845456 Mouse DNA
24	21.8	87.2	180947	9	AL845456 Mouse DNA
25	21.8	87.2	186050	9	BX682537 Mouse DNA
26	21.8	87.2	190928	14	AC118842 CR318639 Mouse DNA
27	21.8	87.2	193326	9	CR318639 Mouse DNA
28	21.8	87.2	196469	14	BX465218 Mus muscu
29	21.8	87.2	198262	14	CR956625 Mus muscu
30	21.8	87.2	200474	14	AC162035 Mus muscu
31	21.8	87.2	201381	9	AL845491 Mouse DNA
32	21.8	87.2	203209	14	AC165368 Mus muscu
33	21.8	87.2	205459	14	CR848808 Mouse DNA
34	21.8	87.2	205573	9	AL591032 Mouse DNA
35	21.8	87.2	206224	14	AC027767 Mus muscu
36	21.8	87.2	207814	9	AL590994 Mouse DNA
37	21.8	87.2	208910	14	AC079573 Mus muscu
38	21.8	87.2	213321	14	AC102988 Rattus no
39	21.8	87.2	216517	9	AL845494 Mouse DNA
40	21.8	87.2	217197	9	BX813319 Mouse DNA
41	21.8	87.2	217867	14	AC095149 Rattus no
42	21.8	87.2	218484	9	AL845476 Mouse DNA
43	21.8	87.2	223248	14	AC055778 Mus muscu
44	21.8	87.2	224765	14	AC106363 Rattus no
45	21.8	87.2	226542	9	BX890623 Mouse DNA
46	21.8	87.2	235381	14	AC128603 Rattus no
47	21.8	87.2	237321	14	AC123015 Rattus no
48	21.8	87.2	238355	14	AC103495 Rattus no
49	21.8	87.2	239591	14	AC128432 Rattus no
50	21.8	87.2	239946	14	AC103505 Rattus no
51	21.8	87.2	241531	14	AC122959 Rattus no
52	21.8	87.2	241795	14	AC123410 Rattus no
53	21.8	87.2	242915	14	AC117981 Rattus no
54	21.8	87.2	245651	14	AC126537 Rattus no
55	21.8	87.2	245993	14	AC130970 Rattus no
56	21.8	87.2	248568	14	AC115184 Rattus no
57	21.8	87.2	251251	14	AC130500 Rattus no
58	21.8	87.2	255563	14	AC125834 Rattus no
59	21.8	87.2	259702	14	AC119031 Rattus no
60	21.8	87.2	274751	14	BX649574 Mus muscu
61	21.8	87.2	278876	14	AC120936 Rattus no
62	20.8	83.2	225269	14	AC133757 Rattus no
63	20.8	83.2	244105	14	AC098897 Rattus no
64	20.2	80.8	878	9	BC017151 Mus muscu
65	20.2	80.8	7279	15	AK120499 Oryza sat
66	20.2	80.8	33961	14	Continuation (6 of AC128246_5
67	20.2	80.8	34046	14	AC094670_5
68	20.2	80.8	40774	14	AC100658 Mus muscu
69	20.2	80.8	47993	14	AC141397 Rattus no
70	20.2	80.8	57402	14	Continuation (4 of AC127888_3
71	20.2	80.8	64074	14	AC117718 Mus muscu
72	20.2	80.8	65102	14	AC120436 Mus muscu
73	20.2	80.8	84257	14	AC099430_3
74	20.2	80.8	106784	14	AC136250 Rattus no
75	20.2	80.8	106784	14	AC136250 Rattus no
76	20.2	80.8	110000	14	Continuation (11 of AC091229_10
77	20.2	80.8	110000	14	AC091242_3
78	20.2	80.8	110000	14	AC091347_1
79	20.2	80.8	110000	14	AC091360 Rattus no
80	20.2	80.8	110000	14	AC094469_1
81	20.2	80.8	110000	14	AC094670_4
82	20.2	80.8	110000	14	AC097996_4
83	20.2	80.8	110000	14	AC106346_1
84	20.2	80.8	110000	14	AC106541_3
85	20.2	80.8	110000	14	AC113631_0
86	20.2	80.8	110000	14	AC114446_3
87	20.2	80.8	110000	14	AC120731_0
88	20.2	80.8	110000	14	AC125991_3
89	20.2	80.8	110000	14	AC127888_2
90	20.2	80.8	110000	14	AC128246_4
91	20.2	80.8	110000	15	AP008215_049

```
92 20.2 80.8 119802 14 AC141020 Rattus no
93 20.2 80.8 125863 14 AC125505 Rattus no
94 20.2 80.8 130254 14 AC136195 Rattus no
c 95 20.2 80.8 141458 14 AC142044 Rattus no
c 96 20.2 80.8 142422 9 AL844588 Mouse DNA
c 97 20.2 80.8 143387 14 AC162885 Mus muscu
c 98 20.2 80.8 146166 14 AC147786 Rattus no
c 99 20.2 80.8 146608 14 AC114719 Rattus no
100 20.2 80.8 147258 14 AC141400 Rattus no
101 20.2 80.8 150849 14 AC141935 Rattus no
102 20.2 80.8 150936 9 AC087722 Rattus no
c 103 20.2 80.8 152788 14 AC120512 Rattus no
c 104 20.2 80.8 153819 9 AC121955 Mus muscu
c 105 20.2 80.8 155094 14 AC122080 Rattus no
c 106 20.2 80.8 157071 9 AC111718 Rattus no
c 107 20.2 80.8 157990 14 AC137311 Rattus no
c 108 20.2 80.8 158089 9 AL645852 Rattus no
c 109 20.2 80.8 160132 14 AC141371 Rattus no
c 110 20.2 80.8 160348 9 AC156275 Mus muscu
c 111 20.2 80.8 160537 9 AC129311 Mus muscu
c 112 20.2 80.8 160901 14 AC120060 Rattus no
c 113 20.2 80.8 161265 9 AL935320 Mouse DNA
c 114 20.2 80.8 163263 14 AC121352 Rattus no
c 115 20.2 80.8 164210 14 AC135478 Rattus no
c 116 20.2 80.8 164571 9 AL929033 Mouse DNA
c 117 20.2 80.8 164826 14 AC124922 Rattus no
c 118 20.2 80.8 165278 14 AC027380 Mus muscu
c 119 20.2 80.8 165316 14 AC119115 Rattus no
c 120 20.2 80.8 167168 15 AP005903 Oryza sat
c 121 20.2 80.8 167526 14 AC121622 Rattus no
c 122 20.2 80.8 167710 14 AC109891 Rattus no
c 123 20.2 80.8 168843 14 AC091711 Rattus no
c 124 20.2 80.8 169294 14 AC112587 Rattus no
c 125 20.2 80.8 169701 14 AC128113 Rattus no
c 126 20.2 80.8 170511 14 AC116263 Rattus no
c 127 20.2 80.8 171292 14 AC118493 Rattus no
c 128 20.2 80.8 171391 9 AC159372 Mus muscu
c 129 20.2 80.8 171817 14 AC027681 Mus muscu
c 130 20.2 80.8 172432 14 AC121221 Rattus no
c 131 20.2 80.8 172786 9 AC069469 Mus muscu
c 132 20.2 80.8 172872 9 AC140456 Mus muscu
c 133 20.2 80.8 173571 14 AC127725 Rattus no
c 134 20.2 80.8 173877 14 AC136010 Rattus no
c 135 20.2 80.8 175586 9 AC117807 Rattus no
c 136 20.2 80.8 175702 14 AC128541 Rattus no
c 137 20.2 80.8 175876 14 AC123336 Rattus no
c 138 20.2 80.8 176602 14 AC095858 Rattus no
c 139 20.2 80.8 177060 14 CR318634 Mus muscu
c 140 20.2 80.8 177227 14 AC117009 Rattus no
c 141 20.2 80.8 177653 9 AC159102 Mus muscu
c 142 20.2 80.8 177885 14 AC141947 Rattus no
c 143 20.2 80.8 178350 9 AC159621 Mus muscu
c 144 20.2 80.8 178671 9 AL805937 Mouse DNA
c 145 20.2 80.8 178815 14 AC127480 Rattus no
c 146 20.2 80.8 179019 9 AC103598 Mus muscu
c 147 20.2 80.8 179341 14 AC128908 Rattus no
c 148 20.2 80.8 179454 14 AC123737 Mus muscu
c 149 20.2 80.8 179481 14 AC034100 Mus muscu
c 150 20.2 80.8 180007 14 AC132933 Mus muscu
```

ALIGNMENTS

```
RESULT 1
AC134563 AC134563 163878 bp DNA linear ROD 19-SEP-2003
LOCUS Mus musculus chromosome 19 clone RP24-179M18, complete sequence.
DEFINITION AC134563
ACCESSION AC134563
VERSION AC134563.5 GI:34850571
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
```

```
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
```

```
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
Wilson,R.K.
The sequence of Mus musculus clone
1 (bases 1 to 163878)
Unpublished
2 (bases 1 to 163878)
McPherson,J.D. and Waterston,R.H.
Direct Submission
Submitted (27-SEP-2002) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
3 (bases 1 to 163878)
Wilson,R.K.
Direct Submission
Submitted (06-SEP-2003) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
4 (bases 1 to 163878)
Wilson,R.K.
Direct Submission
Submitted (19-SEP-2003) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
On Sep 19, 2003 this sequence version replaced gi.34494974.
```

```
----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site:http://genome.wustl.edu
Contact: submissions@watson.wustl.edu
----- Project Information -----
Center project name: M.BB0179M18
-----
```

FEATURES

```
source
1..163878
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="19"
/clone="RP24-179M18"
```

ORIGIN

```
Query Match 93.6%; Score 23.4; DB 9; Length 163878;
Best Local Similarity 96.0%; Pred. No. 1.8;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAAAAGAGCGCAGACTGCTCTCC 25
|||||
Db 47914 AAAAAAGAGCGCAGACTGCTCTCC 47938
```

RESULT 2

```
AC142098/c
LOCUS Mus musculus BAC clone RP24-458F14 from chromosome 19, complete
sequence.
DEFINITION AC142098
ACCESSION AC142098
VERSION AC142098.2 GI:38678651
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
```

```
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 193811)
Veizer,J., Haakenson,W. and Bielicki,L.
The sequence of Mus musculus BAC clone RP24-458F14
Unpublished (2001)
2 (bases 1 to 193811)
Wilson,R.K.
Sequencing of Mus musculus
Unpublished (2001)
3 (bases 1 to 193811)
McPherson,J.D. and Waterston,R.H.
```

```
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
```


TITLE
JOURNAL

REFERENCE
AUTHORS

TITLE
JOURNAL

REFERENCE
AUTHORS

TITLE
JOURNAL

COMMENT

Direct Submission
Submitted (21-MAR-2003) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
4 (bases 1 to 193811)
Wilson.R.K.

Direct Submission
Submitted (04-DEC-2003) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
5 (bases 1 to 193811)
Wilson.R.

Direct Submission
Submitted (01-JAN-2004) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Dec 4, 2003 this sequence version replaced gi:29135708.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu>
Contact: submissions@wustl.wustl.edu
----- Summary Statistics

Center project name: M_BB0458F14

NOTICE: This sequence may not represent the entire insert of this
clone. It may be shorter because we only sequence overlapping
clone sections once, or longer because we provide a small overlap
between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by
restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. Wes Warren,
Department of Genetics, Washington University, St. Louis MO. For
additional information about the map position of this sequence, see
<http://genome.wustl.edu>

SOURCE INFORMATION:

The RPCI-24 BAC Library has been constructed by Pieter de Jong and
coworkers (<http://www.chori.org>) from male C57BL/6J mouse spleen
and/or brain genomic DNA. The clone and detailed information can be
obtained from Pieter de Jong and coworkers at <http://www.chori.org>

NEIGHBORING SEQUENCE INFORMATION:

This sequence is the entire insert of the clone. This clone is
overlapped by AC134563 and AC127271.

FEATURES

source
1. .193811
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="19"
/map="19"
/clone="RP24-458F14"
/clone_lib="RPCI-24"
26944..27849
/note="CpG island (%GC=70.8, o/e=0.74, #CpGs=87)"
62666..62738
/product="tRNA-Ser"
/note="Likely pseudogene (HMM Sc=40.10 / Sec struct
Sc=12.19)"
66009..66081
/product="tRNA-Ser"
/note="Likely pseudogene (HMM Sc=36.29 / Sec struct
Sc=14.73)"
66665..66676
/product="tRNA-Ser"
/note="Likely pseudogene (HMM Sc=33.19 / Sec struct

Sc=-10.52)"
complement(71547..71617)
/product="tRNA-Ser"
/note="Likely pseudogene (HMM Sc=30.48 / Sec struct
Sc=-8.47)"
complement(72364..72436)
/product="tRNA-Ser"
/note="Likely pseudogene (HMM Sc=39.40 / Sec struct
Sc=-17.01)"
75006..75077
/product="tRNA-Ser"
/note="Likely pseudogene (HMM Sc=30.13 / Sec struct
Sc=-3.70)"
complement(76074..76145)
/product="tRNA-Ser"
/note="Likely pseudogene (HMM Sc=35.88 / Sec struct
Sc=-12.12)"
88499..89380
/note="CpG island (%GC=62.5, o/e=0.78, #CpGs=76)"
complement(139660..139731)
/product="tRNA-Ser"
/note="Likely pseudogene (HMM Sc=33.09 / Sec struct
Sc=-11.42)"
complement(173218..173284)
/product="tRNA-Lys"
complement(175362..175434)
/product="tRNA-Ser"
/note="Likely pseudogene (HMM Sc=36.22 / Sec struct
Sc=-14.60)"
185833..185903
/product="tRNA-Ser"
/note="Likely pseudogene (HMM Sc=45.77 / Sec struct
Sc=-12.81)"

ORIGIN

Query Match 93.6%; Score 23.4; DB 9; Length 193811;
Best Local Similarity 96.0%; Pred. No. 1.8;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAAAAGCGCAGACTGCTCTTCC 25
|||||

Db 18517 AAAAAAGCGCAGACTGCTCTTCC 18493

RESULT 3
AC166153
LOCUS
DEFINITION
AC166153.1 GI:71143326
HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
VERSION
KEYWORDS
SOURCE
ORGANISM
Mus musculus
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Murinae; Mus.
1 (bases 1 to 28365)
Birren, B., Nusbaum, C. and Lander, E.
Mus musculus chromosome 7, clone wil-156D7
Unpublished
2 (bases 1 to 28365)
Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N.,
Anderson, M., Anderson, S., Arachchi, H.M., Barna, N., Bastien, V.,
Bloom, T., Boguslavsky, L., Boukhalter, B., Camarata, J., Chang, J.,
Chopel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B.,
DeArellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L.,
Erickson, J., Faro, S., Ferreira, P., FitzGerald, M., Gage, D.,
Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N.,
Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I.,
Johnson, R., Jones, C., Kamat, A., Karatas, A., Kellis, C., Landers, T.,
Levine, R., Lindblad-Toh, K., Liu, G., Liu, X., Lui, A., Mabbitt, R.,
MacLean, C., Macdonald, P., Major, J., Manning, J., Matthews, C.,

McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
 Murphy,T., Naylor,J., Nguyen,C., Nguyen,T., Nicol,R., Norbu,C.,
 O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
 Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C.,
 Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupback,R.,
 Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
 Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J.,
 Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R.,
 Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L.,
 Zimmer,A. and Zody,M.

TITLE JOURNAL

Submitted (26-JUL-2005) Broad Institute of MIT and Harvard, 320
 Charles Street, Cambridge, MA 02141, USA

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Broad Institute of MIT and Harvard

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence_submissions@broad.mit.edu

----- Project Information

Center project name: L33593

Center clone name: 156 D 7

----- Summary Statistics

Sequencing vector: Plasmid; n/a; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 26029 bases at least Q40

Consensus quality: 26780 bases at least Q30

Consensus quality: 27410 bases at least Q20

Insert size: 40000; agarose-fp

Insert size: 27865; sum-of-contigs

Quality coverage: 9.4 in Q20 bases; agarose-fp

Quality coverage: 13.5 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 6 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

* 1
 * 10487: contig of 10487 bp in length
 * 10488 10587: gap of unknown length
 * 10588 11249: contig of 662 bp in length
 * 11250 11349: gap of unknown length
 * 11350 12854: contig of 1505 bp in length
 * 12855 12954: gap of unknown length
 * 12955 14042: contig of 1088 bp in length
 * 14043 14142: gap of unknown length
 * 14143 15428: contig of 1286 bp in length
 * 15429 15528: gap of unknown length
 * 15529 28365: contig of 12837 bp in length.

FEATURES source

1. .28365
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /chromosome="7"
 /map="7"
 /clone="wil-15607"
 /clone_lib="WIBR1 Mouse Fosmid Library"

misc_feature 1. .10487

/note="assembly_fragment
 clone_end:SP6
 vector_side:left"

10488..10587
 /estimated_length=unknown

10588..11249
 /note="assembly_fragment"

11250..11349

misc_feature
 11350..12854
 /note="assembly_fragment"
 gap
 12855..12954
 /estimated_length=unknown
 misc_feature
 12955..14042
 /note="assembly_fragment"
 gap
 14043..14142
 /estimated_length=unknown
 misc_feature
 14143..15428
 /note="assembly_fragment"
 gap
 15429..15528
 /estimated_length=unknown
 misc_feature
 15529..28365
 /note="assembly_fragment"

ORIGIN

Query Match 87.2%; Score 21.8; DB 14; Length 28365;
 Best Local Similarity 92.0%; Pred. No. 11;
 Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAGAGCGCAGACTGGTCTTCC 25

|||||||

Db 12298 AAAAAGAGCACTGACTGGTCTTCC 12322

RESULT 4

CR788287

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Mus musculus

Mus musculus (house mouse)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 98311)

AUTHORS

TITLE

JOURNAL

COMMENT

Submitted (12-FEB-2005) Wellcome Trust Sanger Institute, Hinxton,
 Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
 humquery@sanger.ac.uk
 On Feb 14, 2005 this sequence version replaced gi:58331666.
 Sequence from the Mouse Genome Sequencing Consortium whole genome
 shotgun may have been used to confirm this sequence. Sequence data
 from the whole genome shotgun alone has only been used where it has
 a phred quality of at least 30.

----- Genome Center

Center: Wellcome Trust Sanger Institute

Center code: SC

Web site: http://www.sanger.ac.uk

Contact: humquery@sanger.ac.uk

During sequence assembly data is compared from overlapping clones.
 Where differences are found these are annotated as variations
 together with a note of the overlapping clone name. Note that the
 variation annotation may not be found in the sequence submission
 corresponding to the overlapping clone, as we submit sequences with
 only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all
 regions were either double-stranded or sequenced with an alternate
 chemistry or covered by high quality data (i.e., phred quality >= 30);
 an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by at least
 one plasmid subclone or more than one M13 subclone; and the
 assembly was confirmed by restriction digest, except on the rare
 occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession
 numbers given in the feature table with their source databases:
 Em., EMBL; Swi, SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information

on the WORMPEP database can be found at
http://www.sanger.ac.uk/projects/C_elegans/wormpep/23-362B9 is
from the RP21-23 Mouse BAC Library
constructed by the group of Pieter de Jong.
For further details see http://www.chori.org/bacpac/home.htm
VECTOR: pBACE3.6.

FEATURES

Source
1..98311
Location/Qualifiers
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="2"
/clone="RP23-362B9"
/clone_lib="RP21-23"

ORIGIN

Query Match 87.2%; Score 21.8; DB 9; Length 98311;
Best Local Similarity 92.0%; Pred. No. 12;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAAGGGCAGACTGGCTCTCC 25

Db 39405 AAAAAAGCAGCTGCTGGCTCTCC 39429

RESULT 5

BX546458.1/c

WPCOMMENT

Sequence split into 4 fragments LOCUS BX546458 Accession BX546458

Fragment Name	Begin	End
BX546458.0	1	110000
BX546458.1	100001	210000
BX546458.2	200001	310000
BX546458.3	300001	370067

Continuation (2 of 4) of BX546458 from base 100001 (BX546458 Mus musculus chromosome 2)

Query Match 87.2%; Score 21.8; DB 14; Length 110000;
Best Local Similarity 92.0%; Pred. No. 12;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAAGGGCAGACTGGCTCTCC 25

Db 19233 AAAAAAGCAGCTGCTGGCTCTCC 19209

RESULT 6

BX546462.0

WPCOMMENT

Sequence split into 5 fragments LOCUS BX546462 Accession BX546462

Fragment Name	Begin	End
BX546462.0	1	110000
BX546462.1	100001	210000
BX546462.2	200001	310000
BX546462.3	300001	410000
BX546462.4	400001	447196

LOCUS BX546462 447196 bp DNA linear HTG 24-SEP-2003

DEFINITION Mus musculus chromosome 2 clone RP23-325022, *** SEQUENCING IN PROGRESS ***, 38 unordered pieces.

ACCESSION

BX546462

VERSION

BX546462.3 GI:35209402

KEYWORDS

HTG; HTGS PHASE1.

SOURCE

Mus musculus (house mouse)

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 (bases 1 to 447196)

AUTHORS

Sims, S.

TITLE

Direct Submission

JOURNAL

Submitted (23-SEP-2003) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk

On Sep 24, 2003 this sequence version replaced gi:32563023.

Sequence from the Mouse Genome Sequencing Consortium whole genome shotgun may have been used to confirm this sequence. Sequence data from the whole genome shotgun alone has only been used where it has a phred quality of at least 30.

----- Genome Center

Center: Wellcome Trust Sanger Institute

Center code: SC

Web site: http://www.sanger.ac.uk

Contact: humquery@sanger.ac.uk

----- Project Information

Center project name: BM325022

----- Summary Statistics

Assembly program: XGAP4; version 4.5

Chemistry: Dye-terminator; 100% of reads

Consensus quality: 434558 bases at least Q40

Consensus quality: 438112 bases at least Q30

Consensus quality: 440594 bases at least Q20

Insert size: 443496; sum-of-contigs

Insert size: 189301; 4.9% error; agarose-fp

Quality coverage: 5.28x in Q20 bases; sum-of-contigs Quality

coverage: 14.38x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently consists of 38 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

* 1 14162: contig of 14162 bp in length

* 14163: gap of 100 bp

* 14263: contig of 5397 bp in length

* 19660: gap of 100 bp

* 21781: contig of 2022 bp in length

* 21861: gap of 100 bp

* 21882: contig of 14529 bp in length

* 36411: gap of 100 bp

* 41676: contig of 5166 bp in length

* 41776: gap of 100 bp

* 41777: contig of 12447 bp in length

* 54224: gap of 100 bp

* 54324: contig of 5350 bp in length

* 59674: gap of 100 bp

* 59774: contig of 7398 bp in length

* 67172: gap of 100 bp

* 67272: contig of 2187 bp in length

* 69459: gap of 100 bp

* 69559: contig of 19601 bp in length

* 89160: gap of 100 bp

* 89260: contig of 33372 bp in length

* 122631: contig of 100 bp

* 122732: contig of 11241 bp in length

* 133973: gap of 100 bp

* 134073: contig of 9393 bp in length

* 143466: gap of 100 bp

* 143566: contig of 44052 bp in length

* 187618: gap of 100 bp

* 187718: contig of 80458 bp in length

* 268176: gap of 100 bp

* 308523: contig of 40248 bp in length

* 308524: gap of 100 bp

* 308524: contig of 11639 bp in length

* 320362: gap of 100 bp

* 320363: contig of 3758 bp in length

* 324121: gap of 100 bp

* 324221: contig of 3641 bp in length

* 327862: gap of 100 bp

* 327962: contig of 2499 bp in length

* 330461: gap of 100 bp

* 330560: contig of 3453 bp in length

* 334013: gap of 100 bp

* 334014: contig of 4314 bp in length

* 338427: contig of 4314 bp in length

* 338428 338527: gap of 100 bp
* 338528 341755: contig of 3228 bp in length
* 341756 341855: gap of 100 bp
* 341856 346221: contig of 4366 bp in length
* 346222 346321: gap of 100 bp
* 346322 348507: contig of 2186 bp in length
* 348508 348607: gap of 100 bp
* 348608 378034: contig of 29427 bp in length
* 378035 378134: gap of 100 bp
* 378135 406599: contig of 28625 bp in length
* 406760 406859: gap of 100 bp
* 406860 411480: contig of 4621 bp in length
* 411481 411580: gap of 100 bp
* 411581 414369: contig of 2789 bp in length
* 414370 414469: gap of 100 bp
* 414470 418926: contig of 4457 bp in length
* 418927 419026: gap of 100 bp
* 419027 424085: contig of 5059 bp in length
* 424086 424185: gap of 100 bp
* 424186 426681: contig of 2496 bp in length
* 426682 426781: gap of 100 bp
* 426782 429890: contig of 3109 bp in length
* 429891 429991: contig of 100 bp
* 429991 432321: contig of 2331 bp in length
* 432322 432421: gap of 100 bp
* 432422 435182: contig of 2761 bp in length
* 435183 435282: gap of 100 bp
* 435283 437359: contig of 2077 bp in length
* 437360 437459: gap of 100 bp
* 437460 444723: contig of 7264 bp in length
* 444724 444823: gap of 100 bp
* 444824 447196: contig of 2373 bp in length.

FEATURES

source

Location/Qualifiers

1..447196
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosomes="2"
/clone_lib="RPCI-23"
1..14162
/note="assembly fragment:03333
fragment_chain:1"
14263..19659
/note="assembly fragment:04136
fragment_chain:1"
19760..21781
/note="assembly fragment:01430
fragment_chain:1"
21882..36410
/note="assembly fragment:04198
fragment_chain:1"
36511..41676
/note="assembly fragment:02008
fragment_chain:1"
41777..54223
/note="assembly fragment:04185
fragment_chain:1"
54324..59673
/note="assembly fragment:01187
fragment_chain:1"
59774..67171
/note="assembly fragment:02108
fragment_chain:1"
67272..69458
/note="assembly fragment:04263
fragment_chain:1"
69559..89159
/note="assembly fragment:00095
fragment_chain:2"
89260..122631
/note="assembly fragment:03458
fragment_chain:2"

misc_feature 122732..133972
/note="assembly fragment:02236
fragment_chain:2"
misc_feature 134073..143465
/note="assembly fragment:03351
fragment_chain:3"
misc_feature 143566..187617
/note="assembly fragment:02065
fragment_chain:3"
misc_feature 187718..288175
/note="assembly fragment:03774
fragment_chain:3"
misc_feature 268276..308523
/note="assembly fragment:03651
fragment_chain:4"
misc_feature 308624..320262
/note="assembly fragment:04052
fragment_chain:4"
misc_feature 320363..324120
/note="assembly fragment:00046"
misc_feature 324221..327861
/note="assembly fragment:00226"
misc_feature 327962..330460
/note="assembly fragment:00242"
misc_feature 330561..334013
/note="assembly fragment:00624"
misc_feature 334114..338427
/note="assembly fragment:00667"

Query Match 87.2%; Score 21.8; DB 14; Length 110000;

Best Local Similarity 92.0%; Pred. No. 12;

Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAAGAGCGAGACTGCTCTTC 25

Db 88584 AAAAAAGAGACTGACTGCTCTTC 88608

RESULT 7

CR753898_0/c

WPCOMMENT

Sequence split into 5 fragments LOCUS CR753898 Accession CR753898

Fragment Name	Begin	End
CR753898_0	1	110000
CR753898_1	100001	210000
CR753898_2	200001	310000
CR753898_3	300001	410000
CR753898_4	400001	427662

LOCUS CR753898 427662 bp DNA linear HTG 10-OCT-2004
DEFINITION Mus musculus chromosome 2 clone RP23-100D9, 2 unordered pieces.

ACCESSION CR753898

VERSION CR753898.5 GI:54021828

KEYWORDS HTG; HTGS PHASE1; HTGS CANCELLED.

SOURCE Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 427662)

AUTHORS

Wood, J.

TITLE

JOURNAL

COMMENT

Submitted (09-OCT-2004) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquerry@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Oct 10, 2004 this sequence version replaced gi:53850270.
Sequence from the Mouse Genome Sequencing Consortium whole genome
shotgun may have been used to confirm this sequence. Sequence data
from the whole genome shotgun alone has only been used where it has
a phred quality of at least 30.

----- Genome Center

Center: Wellcome Trust Sanger Institute

Center code: SC

Web site: <http://www.sanger.ac.uk>

Contact: humquery@sanger.ac.uk

----- Project Information

Center project name: BM10009

Assembly program: XGAP4; version 4.5

Chemistry: Dye-terminator; 100% of reads

Consensus quality: 201139 bases at least Q40

Consensus quality: 201243 bases at least Q30

Consensus quality: 204170 bases at least Q20

Insert size: 427562; sum-of-contigs

Insert size: 200670; 4.2% error; agarose-fp

Quality coverage: 3.82x in Q20 bases; sum-of-contigs Quality coverage: 8.39x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently consists of 2 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

* 1 2131: contig of 2131 bp in length

* 2132 2231: gap of 100 bp

* 2232 427662: contig of 425431 bp in length.

Location/Qualifiers

1. .427662

/organism="Mus musculus"

/mol_type="genomic DNA"

/db_xref="taxon:10090"

/chromosome="2"

/clone="RP23-100D9"

/clone_lib="RPCI-23"

misc_feature 1. .2131

/note="assembly_fragment:00090"

2232. .427662

/note="assembly_fragment:02493"

clone end:T7

vector_side:right

ORIGIN

Query Match 87.2%; Score 21.8; DB 14; Length 110000;

Best Local Similarity 92.0%; Pred. No. 12;

Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAAGCGCAGACTGGTCTTCC 25

|||||

DB 53218 AAAAAAGCAGCTGACTGGTCTTCC 53194

|||||

RESULT 8

CR753898 2

WPCOMMENT

Sequence split into 5 fragments LOCUS CR753898 Accession CR753898

Fragment Name	Begin	End
CR753898_0	1	110000
CR753898_1	100001	210000
CR753898_2	200001	310000
CR753898_3	300001	410000
CR753898_4	400001	427662

Continuation (3 of 5) of CR753898 from base 200001 (CR753898 Mus musculus chromosome 2)

Query Match 87.2%; Score 21.8; DB 14; Length 110000;

Best Local Similarity 92.0%; Pred. No. 12;

Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAAGCGCAGACTGGTCTTCC 25

|||||

DB 19431 AAAAAAGCAGCTGACTGGTCTTCC 19455

|||||

RESULT 9

CR753898 2

WPCOMMENT

Contact: humquery@sanger.ac.uk

----- Project Information

Center project name: BM10009

Assembly program: XGAP4; version 4.5

Chemistry: Dye-terminator; 100% of reads

Consensus quality: 201139 bases at least Q40

Consensus quality: 201243 bases at least Q30

Consensus quality: 204170 bases at least Q20

Insert size: 427562; sum-of-contigs

Insert size: 200670; 4.2% error; agarose-fp

Quality coverage: 3.82x in Q20 bases; sum-of-contigs Quality coverage: 8.39x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently consists of 2 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

* 1 2131: contig of 2131 bp in length

* 2132 2231: gap of 100 bp

* 2232 427662: contig of 425431 bp in length.

Location/Qualifiers

1. .427662

/organism="Mus musculus"

/mol_type="genomic DNA"

/db_xref="taxon:10090"

/chromosome="2"

/clone="RP23-100D9"

/clone_lib="RPCI-23"

misc_feature 1. .2131

/note="assembly_fragment:00090"

2232. .427662

/note="assembly_fragment:02493"

clone end:T7

vector_side:right

ORIGIN

Query Match 87.2%; Score 21.8; DB 14; Length 110000;

Best Local Similarity 92.0%; Pred. No. 12;

Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAAGCGCAGACTGGTCTTCC 25

|||||

DB 53218 AAAAAAGCAGCTGACTGGTCTTCC 53194

|||||

RESULT 8

CR753898 2

WPCOMMENT

Sequence split into 5 fragments LOCUS CR753898 Accession CR753898

Fragment Name	Begin	End
CR753898_0	1	110000
CR753898_1	100001	210000
CR753898_2	200001	310000
CR753898_3	300001	410000
CR753898_4	400001	427662

Continuation (3 of 5) of CR753898 from base 200001 (CR753898 Mus musculus chromosome 2)

Query Match 87.2%; Score 21.8; DB 14; Length 110000;

Best Local Similarity 92.0%; Pred. No. 12;

Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAAGCGCAGACTGGTCTTCC 25

|||||

DB 19431 AAAAAAGCAGCTGACTGGTCTTCC 19455

|||||

RESULT 9

CR753898 2

WPCOMMENT

VERSION	CR974476.3	GI:71891486
KEYWORDS	HTG; HTGS PHASE1; HTGS ACTIVEFIN; HTGS_DRAFT; HTGS_FULLTOP.	
SOURCE	Mus musculus (house mouse)	
ORGANISM	Mus musculus	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Mus.	
REFERENCE	1 (bases 1 to 136564)	
AUTHORS	Almeida, J.	
TITLE	Direct Submission	
JOURNAL	Submitted (04-AUG-2005) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vega@sanger.ac.uk	
COMMENT	Clone requests: clonequest@sanger.ac.uk On Aug 5, 2005 this sequence version replaced gi:71833829.	

AC118310	TITLE
LOCUS	JOURNAL
DEFINITION	REFERENCE
ACCESSION	AUTHORS
VERSION	TITLE
KEYWORDS	JOURNAL
SOURCE	REFERENCE
ORGANISM	AUTHORS
REFERENCE	TITLE
AUTHORS	

JOURNAL Submitted (15-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

COMMENT

On Nov 15, 2002 this sequence version replaced gi:23195560. The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GVDU
Center clone name: CH230-373K19
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 136117 bases at least Q40
Consensus quality: 137015 bases at least Q30
Consensus quality: 137564 bases at least Q20
Estimated insert size: 139703; sum-of-contigs estimation
Quality coverage: 8x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently consists of 2 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
* as soon as it is available and the accession number will be preserved.
* 1 123378: contig of 123378 bp in length
* 123379 123478: gap of unknown length
* 123479 143116: contig of 19638 bp in length.

FEATURES

source 1.143116
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-373K19"
/complement(362..1303)
/notes="clone_boundary
clone_end:T7
site:
end sequence: B2237824"
misc_feature 1841..2792
/note="clone_boundary
clone_end:Sp6
site:
end sequence: B2237827"
123379..123478
/estimated_length=unknown
gap

ORIGIN

Query Match 87.2%; Score 21.8; DB 14; Length 143116;
Best Local Similarity 92.0%; Pred. No. 12;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 AAAAAAGCGCGACTGCTTCC 25
|||||
DB 17089 AAAAAAGCGCGACTGCTTCC 17113

RESULT 12
AC166152/c

LOCUS

DEFINITION

AC166152

AC166152

AC166152

AC166152

AC166152

AC166152

AC166152

AC166152

AC166152

AC166152

AC166152

AC166152

AC166152

AC166152

AC166152

AC166152

AC166152

AC166152

AC166152

AC166152

AC166152

AC166152

AC166152

AC166152

AC166152

AC166152

AC166152

AC166152

AC166152

AC166152

AC166152

AC166152

AC166152

AC166152

AC166152

AC166152

AC166152

AC166152

AC166152

AC166152

AC166152

AC166152

AC166152

AC166152

AC166152

AC166152

AC166152

AC166152

AC166152

AC166152

AC166152

AC166152

AC166152

AC166152

AC166152

AC166152

AC166152

AC166152

AC166152

AC166152

AC166152

AC166152

AC166152

AC166152

AC166152

AC166152

AC166152

AC166152

AC166152

AC166152

regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em., EMBL; Sw., SWISSPROT; Tr., TRMBL; Wp., WORMPEP; Information from the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP24-562019 is from a Male (C57BL/6J) mouse BAC library VECTOR: pTARBAC1.

FEATURES
source
1..145110
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="2"
/clone_lib="RPCI-24"
Query Match 87.2%; Score 21.8; DB 9; Length 145110;
Best Local Similarity 92.0%; Pred. No. 12;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

ORIGIN
Query Match 87.2%; Score 21.8; DB 9; Length 145110;
Best Local Similarity 92.0%; Pred. No. 12;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAAGCGGAGCTGCTCTCC 25
AAAAAAAAAGAGCAGTGGTCTCTCC 129908

RESULT 15
CT010482/c
LOCUS
DEFINITION
CT010482.2 GI:72534174
HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
Mus musculus (house mouse)
Mus musculus
Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.
McLaren, S.
Direct Submission
Submitted (10-AUG-2005) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vegas@sanger.ac.uk
Clone requests: clonerequest@sanger.ac.uk
On Aug 12, 2005 this sequence version replaced gi:71142041.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: <http://www.sanger.ac.uk>
Contact: vegas@sanger.ac.uk
----- Project Information
Center project name: bM77N9
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; 100% of reads
Consensus quality: 144894 bases at least Q40
Consensus quality: 145065 bases at least Q30
Consensus quality: 145106 bases at least Q20
Insert size: 145176; sum-of-ctigs
Insert size: 196719; 7.0% error; agarose-fp
Quality coverage: 10.91x in Q20 bases; sum-of-ctigs Quality
coverage: 8.36x in Q20 bases; agarose-fp

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 55513: contig of 55513 bp in length
* 55514 55613: gap of 100 bp
* 55614 142547: contig of 86934 bp in length
* 142548 142647: gap of 100 bp
* 142648 145376: contig of 2729 bp in length.

FEATURES
Location/Qualifiers
1..145376
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="17"
/clone="RP23-77N9"
/clone_lib="RPCI-23"
1..55513
/note="assembly fragment:00038
fragment chain:1
clone end:SP6
vector_side:left"
55614..142547
/note="assembly fragment:00845
fragment chain:1"
142648..145376
/note="assembly fragment:00030"

misc_feature
1..55513
/note="assembly fragment:00038
fragment chain:1
clone end:SP6
vector_side:left"
55614..142547
/note="assembly fragment:00845
fragment chain:1"
142648..145376
/note="assembly fragment:00030"

misc_feature
1..55513
/note="assembly fragment:00038
fragment chain:1
clone end:SP6
vector_side:left"
55614..142547
/note="assembly fragment:00845
fragment chain:1"
142648..145376
/note="assembly fragment:00030"

ORIGIN
Query Match 87.2%; Score 21.8; DB 14; Length 145376;
Best Local Similarity 92.0%; Pred. No. 12;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAAGCGGAGCTGCTCTCC 25
AAAAAAAAAGAGCAGTGGTCTCTCC 85650

RESULT 16
AL607123
LOCUS
DEFINITION
AL607123.22 GI:20338464
HTG.
Mus musculus (house mouse)
Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.
Garner, P.
Direct Submission
Submitted (25-APR-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Apr 29, 2002 this sequence version replaced gi:20068562.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given

in the feature table with their source databases: Em., EMBL; Sw., SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-391E6 is from the RPCI-23 Mouse PAC Library constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: pBACel.6 Location/Qualifiers

```

FEATURES
    source
        1..149369
            /organism="Mus musculus"
            /mol_type="genomic DNA"
            /db_xref="taxon:10090"
            /chromosome="4"
            /clone="RP23-391E6"
            /clone_lib="RPCI-23"

```

ORIGIN

```

Query Match      87.2%; Score 21.8; DB 9; Length 149369;
Best Local Similarity 92.0%; Pred. No. 12;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 AAAAAAGACGCGACACTGGTCTTCC 25
    |||||
Db 19283 AAAAAAGACACAGATTGGTCTTCC 19307
    |||||

```

RESULT 17

```

CR955032/c
LOCUS
DEFINITION
    Homo sapiens clone CHORI-516-BCA274P22, linear HTG 25-MAY-2005
    *** 3 unordered pieces.

```

```

ACCESSION
CR955032.2 GI:66711321
VERSION
HTG; HTGS_PHASE1.
KEYWORDS
SOURCE
    Homo sapiens
    ORGANISM
        Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
        Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
        Homnidae; Homo.

```

REFERENCE

```

AUTHORS
    Sims, S.
TITLE
    Direct Submission
JOURNAL
    Submitted (24-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
    Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
    humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
    On May 25, 2005 this sequence version replaced gi:66392756.

```

COMMENT

```

----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: bCA274P22
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; 100% of reads
Consensus quality: 151132 bases at least Q40
Consensus quality: 151298 bases at least Q30
Consensus quality: 151407 bases at least Q20
Insert size: 151493; sum-of-contigs
Insert size: 187370; 3.1% error; agarose-fp
Quality coverage: 11.97x in Q20 bases; sum-of-contigs Quality
coverage: 9.90x in Q20 bases; agarose-fp
-----

```

* NOTE: This is a 'working draft' sequence. It currently consists of 3 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

```

* 1 56088: contig of 56088 bp in length
* 56089 56188: gap of 100 bp
* 56189 115116: contig of 58928 bp in length
* 115117 115216: gap of 100 bp
* 115217 151693: contig of 36477 bp in length.
FEATURES
    Location/Qualifiers
        1..151693
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            /clone="CHORI-516-BCA274P22"
            /clone_lib="CHORI-516"

```

```

        misc_feature
            1..56088
                /note="assembly_fragment:00606"
                fragment_chain:1
                clone_end:T7
                vector_side:left
                56189..115116
                    /note="assembly_fragment:01538"
                    fragment_chain:1
                    115217..151693
                        /note="assembly_fragment:00047"
                        fragment_chain:1
                        clone_end:SP6
                        vector_side:right

```

```

        misc_feature
            115217..151693
                /note="assembly_fragment:00047"
                fragment_chain:1
                clone_end:SP6
                vector_side:right

```

misc_feature

```

1..56088
    /note="assembly_fragment:00606"
    fragment_chain:1
    clone_end:T7
    vector_side:left
    56189..115116
        /note="assembly_fragment:01538"
        fragment_chain:1
        115217..151693
            /note="assembly_fragment:00047"
            fragment_chain:1
            clone_end:SP6
            vector_side:right

```

ORIGIN

```

Query Match      87.2%; Score 21.8; DB 14; Length 151693;
Best Local Similarity 92.0%; Pred. No. 12;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY

```

1 AAAAAAGACGCGACACTGGTCTTCC 25
    |||||
Db 25889 AAAAAAGACACTGACTGGTCTTCC 25865
    |||||

```

RESULT 18

```

BX511235/c
LOCUS
DEFINITION
    Mouse DNA sequence from clone RP23-232H13 on chromosome 2, complete
    sequence.

```

```

ACCESSION
BX511235
VERSION
BX511235.10 GI:37988139
KEYWORDS
HTG.
SOURCE
    Mus musculus (house mouse)
    ORGANISM
        Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
        Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
        Sciurognathi; Muroidea; Muridae; Murinae; Mus.

```

REFERENCE

```

AUTHORS
    Wood, J.
TITLE
    Direct Submission
JOURNAL
    Submitted (24-OCT-2003) Wellcome Trust Sanger Institute, Hinxton,
    Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
    humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
    On Oct 24, 2003 this sequence version replaced gi:37805699.

```

COMMENT

```

Sequence from the Mouse Genome Sequencing Consortium whole genome
shotgun may have been used to confirm this sequence. Sequence data
from the whole genome shotgun alone has only been used where it has
a phred quality of at least 30.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
-----

```

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above. This sequence was finished as follows unless otherwise noted: all

regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em., EMBL; Sw., SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-232H13 is from the RPCI-23 Mouse BAC Library constructed by the group of Pieter de Jong.

For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: pBACE3.6. Location/Qualifiers

```

FEATURES             Location/Qualifiers
     source            1..169581
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /chromosomes="2"
                     /clone_lib="RPCI-23"

```

ORIGIN

```

Query Match          87.2%; Score 21.8; DB 9; Length 169581;
Best Local Similarity 92.0%; Pred. No. 12;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY      1  AAAAAAGAGCGCAGACTGGTCTTCC 25
        |||||
Db      61942 AAAAAAGAGCAGTGTGCTCTTCC 61918

```

RESULT 19

```

CR955031             174479 bp  DNA  linear  HTG 25-MAY-2005
LOCUS                Homo sapiens clone CHORI-516-BCA362L20, *** SEQUENCING IN PROGRESS
DEFINITION            *** 5 unordered pieces.

```

ACCESSION

```

CR955031             GI:66711320
VERSION              HTG; HTGS PHASE1.
KEYWORDS
SOURCE               Homo sapiens

```

ORGANISM

```

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 174479)
Sims,S.

```

REFERENCE

AUTHORS

TITLE

JOURNAL

```

Direct Submission
Submitted (24-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On May 25, 2005 this sequence version replaced gi:66392755.

```

COMMENT

```

----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: BCA362L20
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; 100% of reads
Consensus quality: 173057 bases at least Q40
Consensus quality: 173562 bases at least Q30
Consensus quality: 173854 bases at least Q20
Insert size: 174079; sum-of-contigs
Insert size: 207259; 2.8% error; agarose-fp
Quality coverage: 6.78x in Q20 bases; sum-of-contigs Quality
coverage: 5.79x in Q20 bases; agarose-fp
-----

```

* NOTE: This is a 'working draft' sequence. It currently consists of 5 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. * This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

```

* 1 80118: contig of 80118 bp in length
* 80119 80218: gap of 100 bp
* 80219 147820: contig of 67602 bp in length
* 147821 147920: gap of 100 bp
* 147921 156970: contig of 9050 bp in length
* 156971 157070: gap of 100 bp
* 157071 165081: contig of 8011 bp in length
* 165082 165181: gap of 100 bp
* 165182 174479: contig of 9298 bp in length.

```

FEATURES

source

```

1..174479
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="CHORI-516-BCA362L20"
/clone_lib="CHORI-516"

```

misc_feature

```

1..80118
/note="assembly_fragment:00975"
fragment_chain:1"
80219..147820
/note="assembly_fragment:00283"
fragment_chain:1"

```

misc_feature

```

147921..156970
/note="assembly_fragment:00186"
fragment_chain:1"
157071..165081
/note="assembly_fragment:00034"
fragment_chain:2"

```

misc_feature

```

165182..174479
/note="assembly_fragment:00104"
fragment_chain:2"
clone_end:T7
vector_side:right"

```

ORIGIN

```

Query Match          87.2%; Score 21.8; DB 14; Length 174479;
Best Local Similarity 92.0%; Pred. No. 12;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY      1  AAAAAAGAGCGCAGACTGGTCTTCC 25
        |||||
Db      119987 AAAAAAGAGCAGTGTGCTCTTCC 120011

```

RESULT 20

CR955031/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

DIRECT SUBMISSION

SUBMITTED

ON MAY 25, 2005

THIS SEQUENCE

VERSION

REPLACED

GI:66392755.

Genome Center

```
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: BCA362L20
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; 100% of reads
Consensus quality: 173057 bases at least Q40
Consensus quality: 173562 bases at least Q30
Consensus quality: 173854 bases at least Q20
Insert size: 174079; sum-of-contigs
Insert size: 207259; 2.8% error; agarose-fp
Quality coverage: 6.78x in Q20 bases; sum-of-contigs Quality
coverage: 5.79x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 5 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 80118: contig of 80118 bp in length
* 80218: gap of 100 bp
* 80219: contig of 67602 bp in length
* 147820: contig of 100 bp
* 147921: contig of 9050 bp in length
* 156970: gap of 100 bp
* 156971: contig of 8011 bp in length
* 157071: contig of 100 bp
* 165081: contig of 8011 bp in length
* 165082: gap of 100 bp
* 165182: contig of 9298 bp in length.
*
* Location/Qualifiers
*   1..174479
*     /organism="Homo sapiens"
*     /mol_type="genomic DNA"
*     /db_xref="taxon:9606"
*     /clone="CHORI-516-BCA362L20"
*     /clone_lib="CHORI-516"
*
* misc_feature
*   1..80118
*     /note="assembly fragment:00975
* fragment_chain:1"
*
* misc_feature
*   80219..147820
*     /note="assembly fragment:00283
* fragment_chain:1"
*
* misc_feature
*   147921..156970
*     /note="assembly fragment:00186
* fragment_chain:1"
*
* misc_feature
*   157071..165081
*     /note="assembly fragment:00034
* fragment_chain:2"
*
* misc_feature
*   165182..174479
*     /note="assembly fragment:00104
* fragment_chain:2
* clone end:T7
* vector_side:right"
*
ORIGIN
Query Match      87.2%; Score 21.8; DB 14; Length 174479;
Best Local Similarity 92.0%; Pred. No. 12;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAAGAGCGCAGACTGGCTCTCC 25
|||||
Db 165208 AAAAAAGAGCAGCTGCTCTTCC 165184

RESULT 21
BX005149 179228 bp DNA linear ROD 18-JUL-2003
LOCUS Mouse DNA sequence from clone RP24-574J13 on chromosome 2, complete
DEFINITION
ACCESSION BX005149
VERSION BX005149
KEYWORDS HTG.

sequence.
BX005149
VERSION BX005149.16 GI:32968213
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.
Tracey, A.
1 (bases 1 to 179228)
Direct Submission
Submitted (18-JUL-2003) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk
Clone requests: clonesrequest@sanger.ac.uk
On July 18, 2003 this sequence version replaced gi:32487195.
Sequence from the Mouse Genome Sequencing Consortium whole genome
shotgun may have been used to confirm this sequence. Sequence data
from the whole genome shotgun alone has only been used where it has
a phred quality of at least 30.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
-----
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest, except on the rare
occasion of the clone being a YAC.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em., EMBL; Swi., SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep/RP24-574J13 is
from a Male (C57BL/6J) mouse BAC Library VECTOR: pIARBA1.
Location/Qualifiers
*   1..179228
*     /organism="Mus musculus"
*     /mol_type="genomic DNA"
*     /db_xref="taxon:10090"
*     /chromosome="2"
*     /clone="RP24-574J13"
*     /clone_lib="RPCI-24"

Query Match      87.2%; Score 21.8; DB 9; Length 179228;
Best Local Similarity 92.0%; Pred. No. 12;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAAGAGCGCAGACTGGCTCTCC 25
|||||
Db 6449 AAAAAAGAGCAGCTGCTCTTCC 6473

RESULT 22
BX005149/c
LOCUS Mouse DNA sequence from clone RP24-574J13 on chromosome 2, complete
DEFINITION
ACCESSION BX005149
VERSION BX005149.16 GI:32968213
KEYWORDS HTG.
```

SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.
1 (bases 1 to 179228)
Tracey A.
Direct Submission
Submitted (18-JUL-2003) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Jul 18, 2003 this sequence version replaced gi:32487195.
Sequence from the Mouse Genome Sequencing Consortium whole genome
shotgun may have been used to confirm this sequence. Sequence data
from the whole genome shotgun alone has only been used where it has
a phred quality of at least 30.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk

During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest, except on the rare
occasion of the clone being a YAC.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em., EMBL; Sw., SWISSPROT; Tr., TrEMBL; Wp., WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C.elegans/wormpep
from the RPI-23 Mouse BAC Library
constructed by the group of Pieter de Jong.
For further details see http://www.chori.org/bacpac/home.htm
VECTOR: pBAC3.6. Location/Qualifiers
1..179228
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="2"
/clone="RP24-574J13"
/clone_lib="RPCI-24"

Query Match 87.2%; Score 21.8; DB 9; Length 179228;
Best Local Similarity 92.0%; Pred. No. 12;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 AAAAAAGAGCGAGCTGGTCTTCC 25
|||||
Db 119388 AAAAAAGAGCGAGCTGGTCTTCC 119364
|||||

RESULT 23
AL845456/c
LOCUS Mouse DNA sequence from clone RP23-90C6 on chromosome 2, complete
DEFINITION
ACCESSION AL845456
VERSION AL845456.10 GI:28445893
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.

REFERENCE
AUTHORS Leongamornlert, D.
TITLE Direct Submission
JOURNAL Submitted (25-FEB-2003) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Feb 20, 2003 this sequence version replaced gi:27817366.
Sequence from the Mouse Genome Sequencing Consortium whole genome
shotgun may have been used to confirm this sequence. Sequence data
from the whole genome shotgun alone has only been used where it has
a phred quality of at least 30.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk

During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest, except on the rare
occasion of the clone being a YAC.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em., EMBL; Sw., SWISSPROT; Tr., TrEMBL; Wp., WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C.elegans/wormpep
from the RPI-23 Mouse BAC Library
constructed by the group of Pieter de Jong.
For further details see http://www.chori.org/bacpac/home.htm
VECTOR: pBAC3.6. Location/Qualifiers
1..180947
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="2"
/clone="RP23-90C6"
/clone_lib="RPCI-23"

Query Match 87.2%; Score 21.8; DB 9; Length 180947;
Best Local Similarity 92.0%; Pred. No. 12;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 AAAAAAGAGCGAGCTGGTCTTCC 25
|||||
Db 142630 AAAAAAGAGCGAGCTGGTCTTCC 142606
|||||

RESULT 24
BX682537
LOCUS Mouse DNA sequence from clone RP23-376N23 on chromosome 2, complete
DEFINITION
ACCESSION BX682537
VERSION BX682537.3 GI:38524372
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.

genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

```

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GVVO
Center clone name: CH230-245G5
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 182925 bases at least Q40
Consensus quality: 183784 bases at least Q30
Consensus quality: 184265 bases at least Q20
Estimated insert size: 188977; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 190928: contig of 190928 bp in length.
FEATURES             source
source               1..190928
                    /organism="Rattus norvegicus"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:10116"
                    /clone="CH230-245G5"
misc_feature         1..1950
                    /note="wgs contig"
misc_feature         2377..3683
                    /note="wgs contig"
misc_feature         182872..183783
                    /note="clone boundary"
                    clone_end:Sp6
                    site:
                    end_sequence:B2140272"
                    184283..186093
                    /note="wgs end extension"
                    clone_end:Sp6"
misc_feature         188925..190928
                    /note="wgs end extension"
                    clone_end:Sp6"

ORIGIN
Query Match          87.2%; Score 21.8; DB 14; Length 190928;
Best Local Similarity 92.0%; Pred. No. 12;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAAGAGCGCAGACTGGTCTTCC 25
    |||||
Db 138712 AAAAAAGAGCAGACTGCTCTTCC 138736

RESULT 26
CR318639 Mouse DNA sequence from clone RP23-317N11 on chromosome 2, complete
LOCUS CR318639 193326 bp DNA linear ROD 18-MAY-2004
DEFINITION Mouse DNA sequence from clone RP23-317N11 on chromosome 2, complete
sequence.
ACCESSION CR318639
VERSION CR318639.6 GI:47496683
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

```

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Mus.

1 (bases 1 to 193326)

Almeida,J.

Direct Submission

Submitted (18-MAY-2004) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk

On May 18, 2004 this sequence version replaced gi:46934930.

Sequence from the Mouse Genome Sequencing Consortium whole genome shotgun may have been used to confirm this sequence. Sequence data from the whole genome shotgun alone has only been used where it has a phred quality of at least 30.

----- Genome Center

Center: Wellcome Trust Sanger Institute

Center code: SC

Web site: http://www.sanger.ac.uk

Contact: humquery@sanger.ac.uk

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WormPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep

from the RPCI-23 Mouse BAC library constructed by the group of Pieter de Jong.

For further details see http://www.chori.org/bacpac/home.htm

VECTOR: pBACe3.6.

FEATURES

source

```

1..193326
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="2"
/clone="RP23-317N11"
/clone_lib="RPCI-23"

```

ORIGIN

Query Match 87.2%; Score 21.8; DB 9; Length 193326;

Best Local Similarity 92.0%; Pred. No. 12;

Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAAGAGCGCAGACTGGTCTTCC 25

Db 178491 AAAAAAGAGCAGCTGCTCTTCC 178515

RESULT 27

BX465218

LOCUS BX465218 196469 bp DNA linear HTG 20-MAY-2003

DEFINITION Mus musculus chromosome 2 clone RP23-419P6, 3 unordered pieces.

ACCESSION BX465218

VERSION BX465218.3 GI:30962340

KEYWORDS HTG; HTGS PHASE1; HTGS_CANCELLED.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

```
REFERENCE
AUTHORS   Sciurognathi; Muroidea; Muridae; Murinae; Mus.
TITLE     1 (bases 1 to 196469)
JOURNAL   Sycamore,N.

COMMENT
Direct Submission
Submitted (18-MAY-2003) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries: clonerequest@sanger.ac.uk
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On May 20, 2003 this sequence version replaced gi:30842655.
Sequence from the Mouse Genome Sequencing Consortium whole genome
shotgun may have been used to confirm this sequence. Sequence data
from the whole genome shotgun alone has only been used where it has
a phred quality of at least 30.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: bm419P6
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; 100% of reads
Consensus quality: 191269 bases at least Q40
Consensus quality: 191438 bases at least Q30
Consensus quality: 191523 bases at least Q20
Insert size: 196269; sum-of-contigs
Insert size: 203083; 4.2% error; agarose-fp
Quality coverage: 6.99x in Q20 bases; sum-of-contigs Quality
coverage: 6.83x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 60380: contig of 60380 bp in length
* 60381 60480: gap of 100 bp
* 60481 72206: contig of 11726 bp in length
* 72207 72306: gap of 100 bp
* 72307 196469: contig of 124163 bp in length.
*
* Location/Qualifiers
* 1..196469
* /organism="Mus musculus"
* /mol_type="genomic DNA"
* /db_xref="taxon:10090"
* /chromosomes="2"
* /clone_lib="RPCI-23"
* /clone="RP23-419P6"
*
* misc_feature
* 1..60380
* /note="assembly_fragment:02020
* clone_end:SP6
* vector_side:left"
*
* misc_feature
* 60481..72206
* /note="assembly_fragment:02063
* fragment_chain:1"
*
* misc_feature
* 72307..196469
* /note="assembly_fragment:02021
* fragment_chain:1"

ORIGIN
Query Match      87.2%; Score 21.8; DB 14; Length 196469;
Best Local Similarity 92.0%; Pred. No. 12;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAAGAGCGAGACTGGTCTTCC 25
    |||||
Db 142575 AAAAAAGAGACTGACTGGTCTTCC 142599

RESULT 28
```

```
CR956625/c
LOCUS     198262 bp      DNA      linear      HTG 10-JUN-2005
DEFINITION Mus musculus chromosome 2 clone RP23-100C10, *** SEQUENCING IN
PROGRESS ***, 13 unordered pieces.
ACCESSION CR956625
VERSION   CR956625.2 GI:67509302
KEYWORDS  HTG; HTGS_PHASE1.
SOURCE    Mus musculus (house mouse)
ORGANISM  Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 198262)
Sims,S.
Direct Submission
Submitted (08-JUN-2005) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vega@sanger.ac.uk
Clone requests: clonerequest@sanger.ac.uk
On Jun 10, 2005 this sequence version replaced gi:66947690.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: vega@sanger.ac.uk
----- Project Information
Center project name: bm100C10
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; 100% of reads
Consensus quality: 193007 bases at least Q40
Consensus quality: 194081 bases at least Q30
Consensus quality: 195094 bases at least Q20
Insert size: 197062; sum-of-contigs
Insert size: 195467; 5.4% error; agarose-fp
Quality coverage: 4.75x in Q20 bases; sum-of-contigs Quality
coverage: 4.82x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 13 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 55953: contig of 55953 bp in length
* 55954 56053: gap of 100 bp
* 56054 63707: contig of 7654 bp in length
* 63708 63807: gap of 100 bp
* 63808 66578: contig of 2771 bp in length
* 66579 66678: gap of 100 bp
* 66679 104450: contig of 37772 bp in length
* 104451 104550: gap of 100 bp
* 104551 116967: contig of 12417 bp in length
* 116968 117067: gap of 100 bp
* 117068 126883: contig of 9816 bp in length
* 126884 126983: gap of 100 bp
* 126984 143917: contig of 16934 bp in length
* 143918 144017: gap of 100 bp
* 144018 147138: contig of 3121 bp in length
* 147139 147238: gap of 100 bp
* 147239 165105: contig of 17867 bp in length
* 165106 165205: gap of 100 bp
* 165206 180021: contig of 14816 bp in length
* 180022 180121: gap of 100 bp
* 180122 187202: contig of 7081 bp in length
* 187203 187302: gap of 100 bp
* 187303 190040: contig of 2738 bp in length
* 190041 190140: gap of 100 bp
* 190141 198262: contig of 8122 bp in length.
*
* Location/Qualifiers
* 1..198262
* /organism="Mus musculus"

FEATURES
source
```


TITLE
JOURNAL
REFERENCE
AUTHORS

```

/misc_feature  /mol_type="genomic DNA"
               /db_xref="taxon:10090"
               /chromosome="2"
               /clone="RP23-100C10"
               /clone_lib="RPC1-23"
               1. .55953
               /notes="assembly fragment:00978"
               fragment chain:1
               clone end:T7
               vector_side:left
               56054. .63707
               /notes="assembly fragment:00094"
               fragment chain:1
               63808. .66578
               /note="assembly fragment:00032"
               fragment chain:1
               66679. .104450
               /notes="assembly fragment:00699"
               fragment chain:1
               104551. .116967
               /notes="assembly fragment:00259"
               fragment chain:1
               117068. .126883
               /notes="assembly fragment:00195"
               fragment chain:1
               126984. .143917
               /notes="assembly fragment:00576"
               fragment chain:1
               144018. .147138
               /notes="assembly fragment:00018"
               fragment chain:1
               147239. .165105
               /notes="assembly fragment:00463"
               fragment chain:1
               165206. .180021
               /notes="assembly fragment:00357"
               fragment chain:1
               180122. .187202
               /notes="assembly fragment:00142"
               fragment chain:1
               187303. .190040
               /notes="assembly fragment:00009"
               fragment chain:1
               clone_end:SP6
               vector_side:right
               190141. .198262
               /notes="assembly fragment:00047"

```

ORIGIN

```

Query Match      87.2%; Score 21.8; DB 14; Length 198262;
Best Local Similarity 92.0%; Pred. No. 12;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY  1  AAAAAAGCGCAGCTGCTTCCTCC 25
    |||||
Db   113064 AAAAAAGCGCAGCTGCTTCCTCC 113040

```

RESULT 29

```

AC162035
LOCUS      AC162035      200474 bp      DNA      linear      HTG 24-JUN-2005
DEFINITION Mus musculus chromosome 18 clone RP23-434N7 map 18, *** SEQUENCING
IN PROGRESS ***
ACCESSION  AC162035
VERSION    AC162035.4 GI:68163656
KEYWORDS   HTG; HTGS_PHRASE2; HTGS_FULLTOP; HTGS_ACTIVEFIN.
SOURCE     Mus musculus
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muridae; Murinae; Mus.
REFERENCE  1 (bases 1 to 200474)
AUTHORS    Birren,B., Nusbaum,C. and Lander,E.

```

Mus musculus chromosome 18, clone RP23-434N7

```

Unpublished
2 (bases 1 to 200474)
Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V.,
Bloom,T., Boguslavsky,L., Boukhgalter,B., Camarata,J., Chang,J.,
Choepe,Y., Collymore,A., Cook,A., Cooke,P., Corum,B.,
DeArellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L.,
Erickson,J., Faro,S., Ferreira,P., Fitzgerald,M., Gage,D.,
Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N.,
Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I.,
Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T.,
Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbitt,R.,
MacLean,C., Macdonald,P., Major,J., Manning,J., Matthews,C.,
McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
Murphy,T., Naylor,J., Nguyen,C., Nguyen,T., Nicol,R., Norbu,C.,
O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C.,
Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R.,
Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zemбек,L.,
Zimmer,A. and Zody,M.
Direct Submission
Submitted (24-MAY-2005) Broad Institute of MIT and Harvard, 320
Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 200474)
Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V.,
Bloom,T., Boguslavsky,L., Boukhgalter,B., Camarata,J., Chang,J.,
Choepe,Y., Collymore,A., Cook,A., Cooke,P., Corum,B.,
DeArellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L.,
Erickson,J., Faro,S., Ferreira,P., Fitzgerald,M., Gage,D.,
Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N.,
Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I.,
Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T.,
Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbitt,R.,
MacLean,C., Macdonald,P., Major,J., Manning,J., Matthews,C.,
McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
Murphy,T., Naylor,J., Nguyen,C., Nguyen,T., Nicol,R., Norbu,C.,
O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C.,
Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R.,
Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zemбек,L.,
Zimmer,A. and Zody,M.
Direct Submission
Submitted (24-JUN-2005) Broad Institute of MIT and Harvard, 320
Charles Street, Cambridge, MA 02141, USA
On Jun 24, 2005 this sequence version replaced gi:67514128.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Broad Institute of MIT and Harvard
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@broad.mit.edu
----- Project Information
Center project name: L32549
Center clone name: 434_N_7
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and

```

* the accession number will be preserved.	
FEATURES source	1 200474: contig of 200474 bp in length.
	Location/Qualifiers
	1. .200474
	/organism="Mus musculus"
	/mol_type="genomic DNA"
ORIGIN	/db_xref="taxon:10090"
	/chromosome="18"
	/map="18"
	/clone="RP23-434N7"
	/clone_lib="RPCI-23 Female Mouse BAC"
Query Match 87.2%; Score 21.8; DB 14; Length 200474;	
Best Local Similarity 92.0%; Pred. No. 12;	
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;	
QY	1 AAAAAAGACGCAGACTGGTCTTCC 25
	152299 AAAAAAGACGACTGCTCTTCC 152323
	Db
RESULT 31	
AC165368	
LOCUS	
DEFINITION	
ACCESSION	
VERSION	
KEYWORDS	
SOURCE	
ORGANISM	
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	
COMMENT	
Center: Wellcome Trust Sanger Institute	
Center code: SC	
Web site: http://www.sanger.ac.uk	
Contact: humquery@sanger.ac.uk	

During sequence assembly data is compared from overlapping clones.	
Where differences are found these are annotated as variations	
together with a note of the overlapping clone name. Note that the	
variation annotation may not be found in the sequence submission	
corresponding to the overlapping clone, as we submit sequences with	
only a small overlap as described above.	
This sequence was finished as follows unless otherwise noted: all	
regions were either double-stranded or sequenced with an alternate	
chemistry or covered by high quality data (i.e., phred quality >=	
30); an attempt was made to resolve all sequencing problems, such	
as compressions and repeats; all regions were covered by at least	
one plasmid subclone or more than one M13 subclone; and the	
assembly was confirmed by restriction digest. The following	
abbreviations are used to associate primary accession numbers given	
in the feature table with their source databases: Em:, EMBL; Sw:,	
SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP	
database can be found at	
http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-360A2 is	
from the RPCI-23 Mouse PAC library	
constructed by the group of Pieter de Jong.	
For further details see http://www.chori.org/bacpac/home.htm	
VECTOR: pBACe3.6.	
FEATURES source	Location/Qualifiers
	1. .201381
	/organism="Mus musculus"
* NOTE: This is a 'working draft' sequence. It currently	
* consists of 49 contigs. The true order of the pieces	
* is not known and their order in this sequence record is	
* arbitrary. Gaps between the contigs are represented as	
* runs of N, but the exact sizes of the gaps are unknown.	
* This record will be updated with the finished sequence	
* as soon as it is available and the accession number will	
* be preserved.	
* 1 1276: contig of 1276 bp in length	
* 1277 1376: gap of unknown length	
* 1377 2699: contig of 1323 bp in length	

2700	2799: gap of unknown length	115446	120862: contig of 5417 bp in length
2800	4360: contig of 1561 bp in length	120863	120962: gap of unknown length
4361	4460: gap of unknown length	120963	126683: contig of 5721 bp in length
4461	5784: contig of 1324 bp in length	126684	126783: gap of unknown length
5785	5884: gap of unknown length	126784	132368: contig of 5585 bp in length
5885	7238: contig of 1354 bp in length	132369	132468: gap of unknown length
7239	7338: gap of unknown length	132469	139665: contig of 7497 bp in length
7339	8584: contig of 1246 bp in length	139666	140065: gap of unknown length
8585	8684: gap of unknown length	140066	147193: contig of 7128 bp in length
8685	10371: contig of 1687 bp in length	147194	147294: gap of unknown length
10372	10471: gap of unknown length	147294	153864: contig of 6571 bp in length
10472	12084: contig of 1613 bp in length	153865	153965: gap of unknown length
12085	12184: gap of unknown length	153966	165095: contig of 11131 bp in length
12185	14087: contig of 1903 bp in length	165096	165195: gap of unknown length
14088	14187: gap of unknown length	165196	176906: contig of 11711 bp in length
14188	15412: contig of 1225 bp in length	176907	177006: gap of unknown length
15413	15512: gap of unknown length	177007	185891: contig of 8885 bp in length
15513	16807: contig of 1295 bp in length	185892	185991: gap of unknown length
16808	16907: gap of unknown length	185992	201588: contig of 15597 bp in length
16909	19393: contig of 2486 bp in length	201589	201688: gap of unknown length
19394	19493: gap of unknown length	201689	203209: contig of 1521 bp in length
19494	21170: contig of 1677 bp in length	FEATURES	
21171	21270: gap of unknown length	Location/Qualifiers	
21271	22666: contig of 1396 bp in length	source	
22667	22766: gap of unknown length	1..203209	
22767	24130: contig of 1364 bp in length	/organism="Mus musculus"	
24131	24230: gap of unknown length	/mol_type="genomic DNA"	
24231	25979: contig of 1749 bp in length	/db_xref="taxon:10090"	
25980	26079: gap of unknown length	/chromosome="1"	
26080	27005: contig of 1626 bp in length	/clone="RP23-28L24"	
27006	27805: gap of unknown length	1..1276	
27806	29900: contig of 2095 bp in length	/note="assembly_name:Contig18"	
29901	30000: gap of unknown length	1277..1376	
30001	3241: contig of 3241 bp in length	/estimated_length=unknown	
3242	3341: gap of unknown length	1377..2699	
3343	36132: contig of 2791 bp in length	/note="assembly_name:Contig19"	
36133	36232: gap of unknown length	2700..2799	
36233	39773: contig of 3541 bp in length	/estimated_length=unknown	
39774	39873: gap of unknown length	2800..4360	
39874	42929: contig of 3056 bp in length	/note="assembly_name:Contig22"	
42930	43029: gap of unknown length	4361..4460	
43030	46035: contig of 3008 bp in length	/estimated_length=unknown	
46036	46135: gap of unknown length	4461..5784	
46136	49576: contig of 3441 bp in length	/note="assembly_name:Contig23"	
49577	49676: gap of unknown length	5785..5884	
49677	52970: contig of 3294 bp in length	/estimated_length=unknown	
52971	53070: gap of unknown length	5885..7238	
53071	56579: contig of 3509 bp in length	/note="assembly_name:Contig24"	
56580	56679: gap of unknown length	7239..7338	
56680	60428: contig of 3749 bp in length	/estimated_length=unknown	
60429	60528: gap of unknown length	7339..8584	
60529	64364: contig of 3836 bp in length	/note="assembly_name:Contig25"	
64365	64464: gap of unknown length	8585..8684	
64465	69850: contig of 5386 bp in length	/estimated_length=unknown	
69851	69950: gap of unknown length	8685..10371	
69951	74459: contig of 4509 bp in length	/note="assembly_name:Contig26"	
74460	74559: gap of unknown length	10372..10471	
74560	78370: contig of 3811 bp in length	/estimated_length=unknown	
78371	78470: gap of unknown length	10472..12084	
78471	83399: contig of 4929 bp in length	/notes="assembly_name:Contig27"	
83400	83499: gap of unknown length	12085..12184	
83500	89000: contig of 5501 bp in length	/estimated_length=unknown	
89001	89100: gap of unknown length	12185..14087	
89101	94291: contig of 5191 bp in length	/note="assembly_name:Contig28"	
94292	94391: gap of unknown length	14088..14187	

```
Query Match      87.2%; Score 21.8; DB 14; Length 203209;
Best Local Similarity 92.0%; Pred. No. 12;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAAGAGCGCAGACTGCTCTTCC 25
   |||||
Db 14973 AAAAAAGAGCAGACTGCTCTTCC 194997

RESULT 32
CR848808      205459 bp      DNA      linear      HTG 03-AUG-2005
LOCUS      Mus musculus chromosome 2 clone RP23-442110, WORKING DRAFT
DEFINITION      SEQUENCE.
ACCESSION      CR848808
VERSION      GI:71794035
KEYWORDS      HTG; HTGS PHASE2; HTGS ACTIVEFIN; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      Mus musculus (house mouse)
ORGANISM      Mus musculus
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
               Sciurognathi; Muridae; Murinae; Mus.
REFERENCE      1 (bases 1 to 205459)
AUTHORS      Tracey A.
TITLE      Direct Submission
JOURNAL      Submitted (02-AUG-2005) Wellcome Trust Sanger Institute, Hinxton,
               Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vegas@sanger.ac.uk
COMMENT      Clone requests: clonerequest@sanger.ac.uk
               On Aug 3, 2005 this sequence version replaced gi:69645812.
               ----- Genome Center
               Center: Wellcome Trust Sanger Institute
               Center code: SC
               Web site: http://www.sanger.ac.uk
               Contact: vegas@sanger.ac.uk
               ----- Project Information
               Center project name: bm442110
               ----- Summary Statistics
               Assembly program: XGAP4; version 4.5
               Chemistry: Dye-terminator; 100% of reads
               Consensus quality: 201342 bases at least Q40
               Consensus quality: 201666 bases at least Q30
               Consensus quality: 201997 bases at least Q20
               Insert size: 205459; sum-of-contigs
               Insert size: 142269; 3.2% error; agarose-fp
               Quality coverage: 11.00x in Q20 bases; sum-of-contigs Quality
               coverage: 16.14x in Q20 bases; agarose-fp
               -----
               * NOTE: This is a 'working draft' sequence. It currently
               * consists of 1 contigs. Gaps between the contigs
               * are represented as runs of N. The order of the pieces
               * is believed to be correct as given, however the sizes
               * of the gaps between them are based on estimates that have
               * been provided by the submitter.
               * This sequence will be replaced
               * by the finished sequence as soon as it is available and
               * the accession number will be preserved.
               *
               1 205459: contig of 205459 bp in length.
               Location/Qualifiers
                 1..205459
                   /organism="Mus musculus"
                   /mol_type="genomic DNA"
                   /db_xref="taxon:10090"
                   /chromosomes="2"
                   /clone="RP23-442110"
                   /clone_lib="RPCI-23"
                   /note="assembly_fragment:01572
                   clone_end:SP6
                   vector_side:left"

misc_feature
1..205459
/note="Tandem repeat. Forced join. Gap size estimated to
be approximately 160bp by restriction digest data."

ORIGIN
Query Match      87.2%; Score 21.8; DB 9; Length 205573;
Best Local Similarity 92.0%; Pred. No. 12;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAAGAGCGCAGACTGCTCTTCC 25
   |||||
Db 189760 AAAAAAGAGCAGACTGCTCTTCC 189784

RESULT 34
```

AC027767/c
LOCUS
DEFINITION

AC027767 206224 bp DNA linear HTG 14-JUN-2003
Mus musculus chromosome 11 clone RP23-51L13 map 11, WORKING DRAFT
SEQUENCE, 5 unordered pieces.

AC027767
VERSION
KEYWORDS
SOURCE
ORGANISM

AC027767.4 GI:31745253
HTG; HTGS PHASE1; HTGS DRAFT; HTGS_FULLTOP.

Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS

1 (bases 1 to 206224)
Birren,B., Nusbaum,C. and Lander,E.

Mus musculus chromosome 11, clone RP23-51L13

Unpublished

2 (bases 1 to 206224)

Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Bada,F.,
Boguslavskiy,L., Boukhgalter,B., Brown,A., Burkett,G.,
Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
Collymore,A., Cooke,P., DeArillano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domino,M., Doyle,M., Ferreira,P., Fitzhugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., LaRoque,K., Lanazares,R., Landers,T., Lehoczy,J.,
Levine,R., Liu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mieng,J., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neil,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
Roy,A., Santos,R., Schauer,S., Severi,P., Spencer,B.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tessaye,S., Theodre,J., Tirrell,A., Travers,M., Trigilio,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,M.

Direct Submission
TITLE
JOURNAL
REFERENCE
AUTHORS

Submitted (01-APR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 206224)

Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
Boguslavskiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y.,
Collymore,A., Cook,A., Cooke,P., Corum,B., DeArillano,K.,
Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,
Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S.,
Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,
Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kellis,C., Landers,T., Levine,R.,
Lindblad-Toh,K., Liu,G., Liu,A., Mabbitt,R., Maclean,C.,
Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
Meldrim,J., Meneus,L., Mihova,T., Mieng,J., Murphy,T., Naylor,J.,
Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
Rachupka,A., Ramasamy,J., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Schauer,S., Schupback,R., Seaman,S., Severi,P., Smith,C.,
Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,
Talamas,J., Tessaye,S., Theodore,J., Topham,K., Travers,M.,
Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission
TITLE
JOURNAL
COMMENT

Submitted (14-JUN-2003) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jun 14, 2003 this sequence version replaced gi:14547878.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WITB
Web site: http://www-seq.wi.mit.edu

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information
Center project name: L8885
Center clone name: 51.L13

----- Summary Statistics
Sequencing vector: M13; M77815; 29% of reads
Sequencing vector: Plasmid; n/a; 71% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 205175 bases at least Q40
Consensus quality: 205524 bases at least Q30
Consensus quality: 205691 bases at least Q20
Insert size: 203000; agarose-fp
Insert size: 205824; sum-of-contigs
Quality coverage: 14.6 in Q20 bases; agarose-fp
Quality coverage: 14.4 in Q20.

* NOTE: This is a 'working draft' sequence. It currently
* consists of 5 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1
* 45102: contig of 45102 bp in length
* 45202: gap of 100 bp
* 45203: 146534: contig of 101332 bp in length
* 146535: 146634: gap of 100 bp
* 146635: 159842: contig of 13208 bp in length
* 159843: 159942: gap of 100 bp
* 159943: 201592: contig of 41650 bp in length
* 201593: 201692: gap of 100 bp
* 201693: 206224: contig of 4532 bp in length.

FEATURES

source	1..206224	Location/Qualifiers
	/organism="Mus musculus"	
	/mol_type="genomic DNA"	
	/db_xref="taxon:10090"	
	/chromosome="11"	
	/map="11"	
	/clone="RP23-51L13"	
	/clone_lib="RPC1-23 Female Mouse BAC"	
misc_feature	1..45102	/notes="assembly_fragment"
	clone_end:SP6	
	vector_side:left	
gap	45103..45202	/estimated_length=100
	45203..146534	/notes="assembly_fragment"
misc_feature	146535..146634	/estimated_length=100
gap	146635..159842	/notes="assembly_fragment"
	159843..159942	/estimated_length=100
misc_feature	159943..201592	/notes="assembly_fragment"
gap	201593..201692	/estimated_length=100
	201693..206224	/notes="assembly_fragment"
misc_feature		clone_end:T7
		vector_side:right

ORIGIN

Query Match 87.2%; Score 21.8; DB 14; Length 206224;
Best Local Similarity 92.0%; Pred. No. 12;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AAAAAAGAGCGGAGCTGCTTCC 25
Db 185703 AAAAAAGAGCGGAGCTGCTTCC 185679

```
RESULT 35
AL590994
LOCUS
DEFINITION
  Mouse DNA sequence from clone RP23-382C19 on chromosome 11,
  complete sequence.
ACCESSION
AL590994
VERSION
AL590994.13 GI:17017757
KEYWORDS
HTG.
SOURCE
Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 207814)
AUTHORS
Lovell J.
TITLE
Direct Submission
JOURNAL
Submitted (17-NOV-2001) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humuery@sanger.ac.uk
COMMENT
On Nov 20, 2001 this sequence version replaced gi:16304740.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em.: EMBL; Sw.:
SWISSPROT; Tr.: TrEMBL; Wp.: WORMPEP; Information on the WORMPEP
database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep
from the RPCI-23 Mouse PAC Library
constructed by the group of Pieter de Jong.
For further details see http://www.chori.org/bacpac/home.htm
VECTOR: pBACE3.6
This sequence is the entire insert of clone RP23-382C19. The true
left end of clone RP23-434F13 is at 189808 in this sequence. The
true right end of clone RP23-328K2 is at 49590 in this sequence.
FEATURES
Source
1..207814
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosomes="11"
/clone="RP23-382C19"
/clone_lib="RPCI-23"
181978
misc_feature
/note="Tandem repeat. Forced join. Gap size estimated to
be approximately 1600bp by restriction digest data."
ORIGIN
Query Match 87.2%; Score 21.8; DB 9; Length 207814;
Best Local Similarity 92.0%; Pred. No. 12;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 AAAAAAGCGCAGACTGGTCTTCC 25
|||||
Db 62730 AAAAAAGCGCAGACTGGTCTTCC 62754
|||||
RESULT 36
AC079573
LOCUS
DEFINITION
  Mus musculus clone RP23-83023, WORKING DRAFT SEQUENCE, 14 unordered
  pieces.
AC079573
VERSION
AC079573.1 GI:9964938
KEYWORDS
HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE
Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 208910)
AUTHORS
DOE Joint Genome Institute.
TITLE
Sequencing of Mouse
JOURNAL
Unpublished
2 (bases 1 to 208910)
AUTHORS
DOE Joint Genome Institute.
TITLE
Direct Submission
JOURNAL
Submitted (02-SEP-2000) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT
Center: Joint Genome Institute
Center Code: JGI
Web site: http://www.jgi.doe.gov
-----
Project Information
Center Project Name: 1769103
Center clone name: RPCI-23_83023
-----
Summary Statistics
Consensus quality: 185662 bases at least Q40
Consensus quality: 198750 bases at least Q30
Consensus quality: 202341 bases at least Q20
Estimated insert size: 233000; agarose-fp estimation
Estimated insert size: 207610; sum-of-contigs estimation
Quality coverage: 6.07 in Q20 bases; agarose-fp estimation
Quality coverage: 6.81 in Q20 bases; sum-of-contigs estimation.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 14 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 1125: contig of 1125 bp in length
* 1126 1225: gap of unknown length
* 1226 2421: contig of 1196 bp in length
* 2422 2521: gap of unknown length
* 2522 4797: contig of 2276 bp in length
* 4798 4897: gap of unknown length
* 4898 6026: contig of 1129 bp in length
* 6027 6126: gap of unknown length
* 6127 7280: contig of 1154 bp in length
* 7281 7380: gap of unknown length
* 7381 10170: contig of 2790 bp in length
* 10171 10270: gap of unknown length
* 10271 26391: contig of 16121 bp in length
* 26392 26491: gap of unknown length
* 26492 44420: contig of 17929 bp in length
* 44421 44520: gap of unknown length
* 44521 66706: contig of 22186 bp in length
* 66707 66806: gap of unknown length
* 66807 95200: contig of 28394 bp in length
* 95201 95300: gap of unknown length
* 95301 124527: contig of 29227 bp in length
* 124528 124627: gap of unknown length
* 124628 156741: contig of 32114 bp in length
* 156742 156841: gap of unknown length
* 156842 187800: contig of 30959 bp in length
* 187801 187900: gap of unknown length
* 187901 208910: contig of 21010 bp in length.
Location/Qualifiers
1..208910
/organism="Mus musculus"
FEATURES
source
```

```

/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="rp23-83023"
/clone_lib="RPCI mouse BAC library 23"
1126..1225
/estimated_length=unknown
2422..2521
/estimated_length=unknown
4798..4897
/estimated_length=unknown
6027..6126
/estimated_length=unknown
7281..7380
/estimated_length=unknown
10171..10270
/estimated_length=unknown
26392..26491
/estimated_length=unknown
44421..44520
/estimated_length=unknown
66707..66806
/estimated_length=unknown
95201..95300
/estimated_length=unknown
124528..124627
/estimated_length=unknown
156742..156841
/estimated_length=unknown
187801..187900
/estimated_length=unknown

ORIGIN
Query Match      87.2%  Score 21.8;  DB 14;  Length 208910;
Best Local Similarity 92.0%;  Pred. No. 12;
Matches 23;  Conservative 0;  Mismatches 2;  Indels 0;  Gaps 0;

QY 1 AAAAAAGAGCGACGACTGTCCTCC 25
Db 148851 AAAAAAGAGCGACGACTGTCCTCC 148875

RESULT 37
AC102988
LOCUS
DEFINITION
Rattus norvegicus clone CH230-60E23, WORKING DRAFT SEQUENCE, 2
unordered pieces.
ACCESSION
AC102988
VERSION
AC102988.5 GI:24819571
KEYWORDS
HTG; HTGS PHASE1; HTGS DRAFT; HTGS_FULLTOP.
SOURCE
Rattus norvegicus (Norway rat)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Rattus.
1 (bases 1 to 213321)
Muzny,D,Marie, Metzker,M, Lee, Abramson,S, Adams,C, Alder,J,
Allen,C, Allen,H, Alsbrooks,S, Amin,A, Angiano,D,
Anyalebechi,V, Aoyagi,A, Ayodeji,M, Baca,E, Baden,H,
Baldwin,D, Bandaranaike,D, Barber,M, Barnstead,M, Benahmed,F,
Blawalo,K, Blair,J, Blankenburg,K, Blyth,P, Brown,M,
Bryant,N, Buhay,C, Burch,P, Burrell,K, Calderon,E,
Cardenas,V, Carter,K, Cavazos,I, Cesar,H, Chen,A,
Chacko,J, Chavez,D, Chen,G, Chen,R, Chen,Y, Chen,Z, Chu,J,
Cleveland,C, Cockrell,R, Cox,C, Coyle,M, Cree,A, D'Souza,L,
Davila,M,L, Davis,C, Davy-Carroll,L, De Anda,C, Dederich,D,
Delgado,O, Denison,S, Deramo,C, Ding,Y, Dinh,H, Divya,K,
Draper,H, Dugan-Rocha,S, Dunn,A, Durbin,K, Duval,B, Eaves,K,
Egan,A, Escotto,M, Eugene,C, Evans,C,A, Falls,T, Fan,G,
Fernandez,S, Finley,M, Flagg,N, Forbes,L, Foster,M, Foster,P,
Fraser,C,M, Gabisi,A, Ganta,R, Garcia,A, Garner,T, Garza,M,
Gebregeorgis,E, Geer,K, Gill,R, Grady,M, Guerra,W, Guevara,W,
Gunaratne,P, Haaland,W, Hamil,C, Hamilton,C, Hamilton,K,
Harvey,Y, Havlak,P, Hawes,A, Henderson,N, Hernandez,J,
Hernandez,R, Hines,S, Hladun,S,L, Hodgson,A, Hognes,M,
Hollins,B, Howells,S, Hulyk,S, Hume,J, Idlebird,D, Jackson,A,
Jackson,L, Jacob,S, Jiang,H, Johnson,B, Johnson,R, Jollivet,A,
Karpathy,S, Kelly,S, Khan,Z, King,L, Kovar,C,
Kowis,C, Kraft,C,L, Lebow,H, Levan,J, Lewis,L, Li,Z, Liu,J,
Liu,J, Liu,W, Liu,Y, London,P, Longacre,S, Lopez,J,
Lorensuewa,L, Loulseged,H, Lorado,R,J, Lu,X, Ma,J,
Maheshwari,M, Mahindartne,M, Mahmoud,M, Malloy,K, Mangum,A,
Mangum,B, Mapua,P, Martin,K, Martin,R, Martinez,E,
Mahoney,S, McLeod,M,P, McNeill,T,Z, Meenen,E,
Milosavljevic,A, Miner,G, Minja,E, Montemayor,J, Moore,S,
Morgan,M, Morris,K, Morris,S, Munidasa,M, Murphy,M, Nair,L,
Nankervis,C, Neal,D, Newton,N, Nguyen,N, Norris,S, Parks,K,
Nwaokemele,O, Okwuonu,G, Olarnpunsagoon,A, Pal,S, Parks,K,
Pasternak,S, Paul,H, Perez,A, Perez,L, Pfannkuch,C,
Plopper,F, Poindexter,A, Popovic,D, Primus,E, Pu,L, L,
Puazo,M, Quiroz,J, Rachlin,E, Reeves,K, Regier,M,A, Reigh,R,
Reilly,B, Reilly,M, Ren,Y, Reuter,M, Richards,S, Riggs,F,
Rives,C, Rodkey,T, Rojas,A, Rose,M, Rose,R, Ruiz,S,J,
Sanders,W, Savary,G, Scherer,S, Scott,G, Shatsman,S, Shen,H,
Shetty,J, Shvartsbeyn,A, Sisson,I, Sitter,C,D, Smajls,D,
Steimle,M, Strong,R, Sutton,A, Svatek,A, Tabor,P, Taylor,C,
Taylor,T, Thomas,N, Thomas,S, Tingey,A, Trejos,Z, Usmani,K,
Valas,R, Vera,V, Villasana,D, Waldron,L, Walker,B, Wang,J,
Wang,Q, Wang,S, Warren,J, Warren,R, Wei,X, White,F,
Williams,G, Willson,R, Wlaczek,R, Wooden,H, Worley,K,
Wright,D, Wright,R, Wu,J, Yakub,S, Yen,J, Yoon,L, Yoon,V,
Yu,F, Zhang,J, Zhou,J, Zhou,X, Zhao,S, Dunn,D, von
Niederhausern,A, Weiss,R, Smith,D,R, Holt,R,A, Smith,H,O,
Weinstock,G, and Gibbs,R,A.
Direct Submission
Unpublished
2 (bases 1 to 213321)
Worley,K,C
Direct Submission
Submitted (26-NOV-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 213321)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (09-NOV-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Nov 9, 2002 this sequence version replaced gi:23101319.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GHGS
Center clone name: CH230-60E23
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 203322 bases at least Q40
Consensus quality: 205366 bases at least Q30
Consensus quality: 206688 bases at least Q20
Estimated insert size: 211097; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

```

 * NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 2 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

* 1 112868: contig of 112868 bp in length
 * 112869 112968: gap of unknown length
 * 112969 213321: contig of 100353 bp in length.

FEATURES

source

1. .213321
 Location/Qualifiers
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-60E23"
 complement(95870..96694)
 /note="clone boundary
 clone end:Sp6
 site:EcoRI
 end_sequence:BH364717"

misc_feature

112869..112968
 /estimated_length=unknown
 211383..213321
 /note="wgs_contig"

gap

211383..213321
 /note="wgs_contig"

misc_feature

ORIGIN

Query Match 87.2%; Score 21.8; DB 14; Length 213321;
 Best Local Similarity 92.0%; Pred. No. 12;
 Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAAGCGCAGACTGCTCTCC 25
 |||||
 Db 8548 AAAAAAGCAGCAGACTGCTCTCC 8572

RESULT 38

AL845494

LOCUS

DEFINITION Mouse DNA sequence from clone RP23-330D3 on chromosome 2, complete
 sequence.

ACCESSION

AL845494

VERSION

AL845494.3

KEYWORDS

HTG.

SOURCE

Mus musculus

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muridea; Muridae; Murinae; Mus.

REFERENCE

Tracey, A.

AUTHORS

Direct Submission

TITLE

Submitted (09-OCT-2002) Wellcome Trust Sanger Institute, Hinxton,
 Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
 humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk

JOURNAL

On Oct 12, 2002 this sequence version replaced gi:23337886.

COMMENT

----- Genome Center
 Center: Wellcome Trust Sanger Institute
 Center code: SC
 Web site: <http://www.sanger.ac.uk>
 Contact: humquery@sanger.ac.uk

 During sequence assembly data is compared from overlapping clones.
 Where differences are found these are annotated as variations
 together with a note of the overlapping clone name. Note that the
 variation annotation may not be found in the sequence submission
 corresponding to the overlapping clone, as we submit sequences with
 only a small overlap as described above.
 This sequence was finished as follows unless otherwise noted: all

regions were either double-stranded or sequenced with an alternate
 chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by at least
 one plasmid subclone or more than one M13 subclone; and the
 assembly was confirmed by restriction digest. The following
 abbreviations are used to associate primary accession numbers given
 in the feature table with their source databases: Em., EMBL; Swi.,
 SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information on the WORMPEP
 database can be found at
http://www.sanger.ac.uk/Projects/c_elegans/wormpep RP23-330D3 is
 from the RPCI-23 Mouse PAC Library
 constructed by the group of Pieter de Jong.
 For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: pBAC3.6.
 Location/Qualifiers
 1. .216517
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /chromosome="2"
 /clone="RP23-330D3"
 /clone_lib="RPCI-23"

FEATURES

source

1. .216517
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /chromosome="2"
 /clone="RP23-330D3"
 /clone_lib="RPCI-23"

ORIGIN

Query Match 87.2%; Score 21.8; DB 9; Length 216517;
 Best Local Similarity 92.0%; Pred. No. 12;
 Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAAGCGCAGACTGCTCTCC 25
 |||||
 Db 186644 AAAAAAGCAGCTGCTCTCC 186668

RESULT 39

BX813319/c

LOCUS

DEFINITION Mouse DNA sequence from clone RP23-54B12 on chromosome 2, complete
 sequence.

ACCESSION

BX813319

VERSION

BX813319.4

KEYWORDS

HTG.

SOURCE

Mus musculus

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muridea; Muridae; Murinae; Mus.

REFERENCE

Sycamore, N.

AUTHORS

Direct Submission

TITLE

Submitted (06-FEB-2004) Wellcome Trust Sanger Institute, Hinxton,
 Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
 humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk

JOURNAL

On Feb 5, 2004 this sequence version replaced gi:41060267.

COMMENT

Sequence from the Mouse Genome Sequencing Consortium whole genome
 shotgun may have been used to confirm this sequence. Sequence data
 from the whole genome shotgun alone has only been used where it has
 a phred quality of at least 30.

----- Genome Center

Center: Wellcome Trust Sanger Institute

Center code: SC

Web site: <http://www.sanger.ac.uk>Contact: humquery@sanger.ac.uk

 During sequence assembly data is compared from overlapping clones.
 Where differences are found these are annotated as variations
 together with a note of the overlapping clone name. Note that the
 variation annotation may not be found in the sequence submission
 corresponding to the overlapping clone, as we submit sequences with
 only a small overlap as described above.
 This sequence was finished as follows unless otherwise noted: all
 regions were either double-stranded or sequenced with an alternate
 chemistry or covered by high quality data (i.e., phred quality >=

30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em; EMBL; Swi; SWISSPROT; Tr; TREMBL; Wp; WORMPEP; Information on the WORMPEP database can be found at

http://www.sanger.ac.uk/projects/C_elegans/wormpep RP23-54B12 is from the RPCI-23 Mouse BAC library

constructed by the group of Pieter de Jong.
For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: pBAC3.6

FEATURES

source

```
1. 217197 Location/Qualifiers
    /organism="Mus musculus"
    /mol_type="genomic DNA"
    /db_xref="taxon:10090"
    /chromosome="2"
    /clone="RP23-54B12"
    /clone_lib="RPCI-23"
```

ORIGIN

Query Match 87.2%; Score 21.8; DB 9; Length 217197;

Best Local Similarity 92.0%; Pred. No. 12;

Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAAGCGCAGACTGCTTCC 25

||||| | | | | | | | | | | | | | | | | | | | |

Db 91248 AAAAAAGCGCAGCTGCTTCC 91224

RESULT 40

AC095149

LOCUS

DEFINITION Rattus norvegicus clone CH230-813, WORKING DRAFT SEQUENCE, 3 linear HTG 09-MAY-2003

unordered pieces.

AC095149 GI:30467261

HTG; HTGS PHASE1; HTGS DRAFT; HTGS_FULLTOP.

Rattus norvegicus (Norway rat)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

Sciurognathi; Muridae; Murinae; Rattus.

1 (bases 1 to 217867)

Muzny,D,Marie, Metzker,M, Lee, A., Abramson, S., Amin, A., Anguiano, D.,

Allen, C., Allen, H., Alsbrooks, S., Ayodeji, M., Baca, E., Baden, H.,

Anyalebechi, V., Aoyagi, A., Ayodeji, M., Barber, M., Barnstead, M., Benahmed, F.,

Baldwin, D., Bandaranaike, D., Barch, P., Burrell, K., Calderon, E.,

Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,

Bryant, N., Buhay, C., Burch, P., Burrell, K., Ceasar, H., Center, A.,

Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Chen, Z., Chu, J.,

Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,

Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,

Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,

Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,

Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,

Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G.,

Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,

Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,

Gebrgeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,

Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,

Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,

Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M.,

Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,

Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,

Karpachy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,

Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,

Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,

Lorensuhewa, L., Loulseghe, H., Lozada, R.J., Lu, X., Ma, J.,

Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A.,

Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
Mawhney, S., McLeod, M.P., McNeill, T.Z., Meenen, E.,
Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
Morgan, M., Morris, K., Morris, S., Munitasa, M., Murphy, M., Nair, L.,
Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
Nwaokemele, O., Okwunnu, G., Olarunpasa, A., Pal, S., Parks, K.,
Pascernak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,
Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L., L.,
Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R.,
Reilly, B., Reilly, M., Reilly, X., Reuter, M., Richards, S., Riggs, F.,
Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J.,
Sanders, W., Savary, G., Scherer, S., Scott, G., Shatman, S., Shen, H.,
Shetty, J., Shivabeyn, A., Sisson, I., Sitter, C.D., Smajic, D.,
Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,
Steimle, M., Strong, R., Sutton, A., Svatek, K., Tabor, P., Taylor, C.,
Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Uman, K.,
Valas, R., Vera, V., Villanasa, D., Waldron, L., Walker, B., Wang, J.,
Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, P.,
Williams, G., Willson, R., Wleczek, R., Wooden, H., Worley, K.,
Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,
Weinstock, G. and Gibbs, R.A.

Direct Submission

Unpublished

2 (bases 1 to 217867)

Worley, K.C.

Direct Submission

JOURNAL

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Submitted (16-SEP-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 217867)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (09-MAY-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On May 9, 2003 this sequence version replaced gi:24940815.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.

----- Genome Center -----

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information -----

Center project name: GCTH

Center clone name: CH230-813

----- Summary Statistics -----

Assembly program: Atlas;

Consensus quality: 205299 bases at least Q40

Consensus quality: 207638 bases at least Q30

Consensus quality: 208930 bases at least Q20

Estimated insert size: 216123; sum-of-contigs estimation

Quality coverage: 8x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).

* NOTE: This is a 'working draft' sequence. It currently
consists of 3 contigs. The true order of the pieces

is not known and their order in this sequence record is

arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

* 1 65534: contig of 65534 bp in length
 * 65535 65634: gap of unknown length
 * 65635 182468: contig of 116834 bp in length
 * 182469 182568: gap of unknown length
 * 182569 217867: contig of 35299 bp in length.

FEATURES
 source
 1. .217867
 Location/Qualifiers
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-813"
 65535..65634
 /estimated_length=unknown
 182469..182568
 /estimated_length=unknown

ORIGIN

Query Match 87.2%; Score 21.8; DB 14; Length 217867;
 Best Local Similarity 92.0%; Pred. No. 12;
 Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 AAAAAAGCGCAGACTGGTCTTCC 25
 ||||| ||||| ||||| ||||| |||||
 Db 17901 AAAAAAGCGCAGCTGCTCTTCC 17925

RESULT 41
 AL845476/c 218484 bp DNA linear ROD 15-OCT-2003
 LOCUS
 DEFINITION Mouse DNA sequence from clone RP23-355G16 on chromosome 2, complete sequence.
 ACCESSION AL845476
 VERSION AL845476.15 GI:37693439
 KEYWORDS HTG.
 SOURCE Mus musculus (house mouse)
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 218484)
 Almeida,J.

REFERENCE
 AUTHORS
 TITLE Direct Submission
 JOURNAL Submitted (15-OCT-2003) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
 COMMENT On Oct 15, 2003 this sequence version replaced gi:32452825. Sequence from the Mouse Genome Sequencing Consortium whole genome shotgun may have been used to confirm this sequence. Sequence data from the whole genome shotgun alone has only been used where it has a phred quality of at least 30.
 ----- Genome Center
 Center: Wellcome Trust Sanger Institute
 Center code: SC
 Web site: <http://www.sanger.ac.uk>
 Contact: humquery@sanger.ac.uk

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
 This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:
 Em., EMBL; Sw., SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-355G16 is from the RPCI-23 Mouse BAC Library constructed by the group of Pieter de Jong.
 For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: pBAC3.6.
 Location/Qualifiers

1. .218484
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /chromosome="2"
 /clone="RP23-355G16"
 /clone_lib="RPCI-23"

ORIGIN

Query Match 87.2%; Score 21.8; DB 9; Length 218484;
 Best Local Similarity 92.0%; Pred. No. 12;
 Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 AAAAAAGCGCAGACTGGTCTTCC 25
 ||||| ||||| ||||| ||||| |||||
 Db 113311 AAAAAAGCAGCTGCTCTTCC 113287

RESULT 42
 AC055778 223248 bp DNA linear HTG 22-MAR-2003
 LOCUS
 DEFINITION Mus musculus chromosome 10 clone RP23-76A3 map 10, WORKING DRAFT SEQUENCE, 12 unordered pieces.
 ACCESSION AC055778
 VERSION AC055778.4 GI:29150442
 KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
 SOURCE Mus musculus (house mouse)
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
 AUTHORS Birren,B., Nusbaum,C. and Lander,E.
 TITLE Mus musculus chromosome 10, clone RP23-76A3
 JOURNAL Unpublished
 REFERENCE
 AUTHORS 2 (bases 1 to 223248)
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N., Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F., Boguslavskiy,L., Boukhgalter,B., Brown,A., Burkett,G., Campiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S., Collumore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Domino,M., Doyle,M., Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L., Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J., Lacroque,K., Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N., McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R., Meldrum,J., Meneus,I., Mihova,T., Miranda,C., Mlenga,V., Morrow,J., Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N., Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Severi,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J., Testafaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zimmer,A. and Zody,M.

Direct Submission
 Submitted (18-APR-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 3 (bases 1 to 223248)
 Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,

Boguslavskiy, L., Boukhgalter, B., Camarata, J., Chang, J., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B., DeArellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagoopian, D., Hagos, B., Hall, J., Horton, L., Hulme, M., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Lui, A., Mabbitt, R., MacLean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Melgrim, J., Meneus, L., Mihova, P., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramaamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schuback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (22-MAR-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 22, 2003 this sequence version replaced gi:13357446.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

TITLE JOURNAL

COMMENT

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L5749

Center clone name: 76_A_3

----- Summary Statistics

Sequencing vector: M13; M7815; 28% of reads

Sequencing vector: Plasmid; n/a; 72% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 220412 bases at least Q40

Consensus quality: 221187 bases at least Q30

Consensus quality: 221711 bases at least Q20

Insert size: 204000; agarose-fp

Insert size: 222148; sum-of-contents

Quality coverage: 13.2 in Q20 bases; agarose-fp

Quality coverage: 12.1 in Q20 b.

* NOTE: This is a 'working draft' sequence. It currently
* consists of 12 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 28669; contig of 28669 bp in length
* 28670 28769; gap of 100 bp
* 28770 28745; contig of 976 bp in length
* 29746 29845; gap of 100 bp
* 29846 32902; contig of 3057 bp in length
* 32903 33002; gap of 100 bp
* 33003 40651; contig of 7649 bp in length
* 40652 40751; gap of 100 bp
* 40752 49207; contig of 8456 bp in length
* 49208 49308; gap of 100 bp
* 49308 57881; contig of 8474 bp in length
* 57882 57881; gap of 100 bp
* 57882 57875; contig of 10794 bp in length
* 57876 68775; gap of 100 bp
* 68776 84586; contig of 15811 bp in length
* 84587 84686; gap of 100 bp
* 84687 101123; contig of 16437 bp in length
* 101124 101224; gap of 100 bp
* 101224 122063; contig of 20840 bp in length
* 122064 122163; gap of 100 bp

* 122164 153291; contig of 31128 bp in length
* 153292 153391; gap of 100 bp
* 153392 223248; contig of 69857 bp in length.

FEATURES
source
1..223248
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="10"
/map="10"
/clone="RP23-76A3"
/clone_lib="RPCI-23 Female Mouse BAC"
1..28669
/note="assembly_fragment"
clone_end:SP6
vector_side:left
28670..28769
/estimated_length=100
28770..29745
/note="assembly_fragment"
29746..29845
/estimated_length=100
29846..32902
/note="assembly_fragment"
32903..33002
/estimated_length=100
33003..40651
/note="assembly_fragment"
40652..40751
/estimated_length=100
40752..49207
/note="assembly_fragment"
49208..49307
/estimated_length=100
49308..57781
/note="assembly_fragment"
57782..57881
/estimated_length=100
57882..68675
/note="assembly_fragment"
68676..68775
/estimated_length=100
68776..84586
/note="assembly_fragment"
84587..84686
/estimated_length=100
84687..101123
/note="assembly_fragment"
101124..101223
/estimated_length=100
101224..122063
/note="assembly_fragment"
122064..122163
/estimated_length=100
122164..153291
/note="assembly_fragment"
153292..153391
/estimated_length=100
153392..223248
/note="assembly_fragment"
clone_end:T7
vector_side:right

ORIGIN

Query Match 87.2%; Score 21.8; DB 14; Length 223248;
Best Local Similarity 92.0%; Pred. No. 12;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AAAAAAGCGCAGACTGGCTTCC 25

Db 77657 AAAAAAGCAGATTGGCTTCC 77681

RESULT 44
 BX890623
 LOCUS
 DEFINITION
 Mouse DNA sequence from clone RP23-83023 on chromosome 2, complete sequence.

ACCESSION
 BX890623
 VERSION
 BX890623.6 GI:43823581
 KEYWORDS
 HTG.

SOURCE
 Mus musculus (house mouse)

ORGANISM
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 226542)

REFERENCE
 AUTHORS
 TITLE
 JOURNAL

COMMENT
 Direct Submission
 Submitted (26-FEB-2004) Wellcome Trust Sanger Institute, Hinxton,
 Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
 humquery@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk
 On Feb 26, 2004 this sequence version replaced gi:42627462.
 Sequence from the Mouse Genome Sequencing Consortium whole genome
 shotgun may have been used to confirm this sequence. Sequence data
 from the whole genome shotgun alone has only been used where it has
 a phred quality of at least 30.
 ----- Genome Center
 Center: Wellcome Trust Sanger Institute
 Center code: SC
 Web site: <http://www.sanger.ac.uk>
 Contact: humquery@sanger.ac.uk

During sequence assembly data is compared from overlapping clones.
 Where differences are found these are annotated as variations
 together with a note of the overlapping clone name. Note that the
 variation annotation may not be found in the sequence submission
 corresponding to the overlapping clone, as we submit sequences with
 only a small overlap as described above.
 This sequence was finished as follows unless otherwise noted: all
 regions were either double-stranded or sequenced with an alternate
 chemistry or covered by high quality data (i.e. phred quality >=
 30); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by at least
 one plasmid subclone or more than one M13 subclone; and the
 assembly was confirmed by restriction digest, except on the rare
 occasion of the clone being a YAC.
 The following abbreviations are used to associate primary accession
 numbers given in the feature table with their source databases:
 Em., EMBL; Sw., SWISSPROT; Tr., TrEMBL; Wp., WORMPEP; Information
 on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-83023 is
 from the RPCI-23 Mouse BAC Library
 constructed by the group of Pieter de Jong.
 For further details see <http://www.chori.org/bacpac/home.htm>
 VECTOR: pBACe3.6.

FEATURES
 source
 Location/Qualifiers
 1..226542
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /chromosome="2"
 /clone="RP23-83023"
 /clone_lib="RPCI-23"

ORIGIN
 Query Match 87.2%; Score 21.8; DB 9; Length 226542;
 Best Local Similarity 92.0%; Pred. No. 12;
 Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AAAAAAGCCGACACTGGCTTCC 25
 |||||
 Db 41228 AAAAAAGCCGACACTGGCTTCC 41252
 |||||

RESULT 45

AC128603/c
 LOCUS
 DEFINITION
 Rattus norvegicus clone CH230-227D15, *** SEQUENCING IN PROGRESS
 ***. 4 unordered pieces.

ACCESSION
 AC128603
 VERSION
 AC128603.2 GI:23196137
 KEYWORDS
 HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
 SOURCE
 Rattus norvegicus (Norway rat)

ORGANISM
 Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muridae; Murinae; Rattus.
 1 (bases 1 to 229093)

REFERENCE
 AUTHORS
 TITLE
 JOURNAL

COMMENT
 Direct Submission
 Submitted (19-JUL-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 229093)
 Rat Genome Sequencing Consortium.
 Direct Submission

AC128603
 LOCUS
 DEFINITION
 Rattus norvegicus clone CH230-227D15, *** SEQUENCING IN PROGRESS
 ***. 4 unordered pieces.

ACCESSION
 AC128603
 VERSION
 AC128603.2 GI:23196137
 KEYWORDS
 HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
 SOURCE
 Rattus norvegicus (Norway rat)

ORGANISM
 Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muridae; Murinae; Rattus.
 1 (bases 1 to 229093)

REFERENCE
 AUTHORS
 TITLE
 JOURNAL

COMMENT
 Direct Submission
 Submitted (19-JUL-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 229093)
 Rat Genome Sequencing Consortium.
 Direct Submission

JOURNAL Submitted (19-SEP-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

COMMENT On Sep 19, 2002 this sequence version replaced gi:21909406. The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). As a result, the sequence may extend beyond the ends of the clone and there may be contigs that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GYMO
Center clone name: CH230-227D15
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 214373 bases at least Q40
Consensus quality: 216710 bases at least Q30
Consensus quality: 218218 bases at least Q20
Estimated insert size: 236256; sum-of-contigs estimation
Quality coverage: 4x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently consists of 4 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

FEATURES
source
1 135103: contig of 135103 bp in length
* 135104: gap of unknown length
* 135204 176769: contig of 41566 bp in length
* 176770 176869: gap of unknown length
* 176870 182770: contig of 5901 bp in length
* 182771 182870: gap of unknown length
* 182871 229093: contig of 46223 bp in length.

misc_feature
44238..45703
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-227D15"
/notes="wgs contig"

gap
135104..135203
/estimated_length=unknown
135204..136476
/notes="wgs contig"

misc_feature
176770..176869
/estimated_length=unknown
182771..182870
/estimated_length=unknown

ORIGIN
Query Match 87.2%; Score 21.8; DB 14; Length 229093;
Best Local Similarity 92.0%; Pred. No. 12;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAAGCGCGACTGCTTCC 25
|||||
Db 132333 AAAAAAGCGCGACTGCTTCC 132309
|||||

RESULT 46
AC123015

LOCUS AC123015 235381 bp DNA linear HTG 15-NOV-2002
DEFINITION Rattus norvegicus clone CH230-154H6, *** SEQUENCING IN PROGRESS
***, 3 unordered pieces.
AC123015
AC123015.4 GI:25007936
HTG: HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
KEYWORDS Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Rattus;
1 (bases 1 to 235381)
Muzny,D,Marie., Metzker,M, Lee., Abramzon,S., Adams,C., Alder,J., Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D., Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H., Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F., Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M., Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E., Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J., Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L., Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D., Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K., Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K., Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G., Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P., Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M., Gebregorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,M., Gunaratne,P., Haaland,W., Hamill,C., Hamilton,C., Hamilton,K., Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J., Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M., Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A., Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A., Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C., Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J., Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J., Lorensuewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J., Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A., Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E., Mawhinney,S., McLeod,M.P., McNeill,T.Z., Meenen,E., Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S., Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L., Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Nwackelameh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K., Pasternak,S., Paul,H., Perez,A., Perez,L., Pfankoch,C., Plopper,F., Poindexter,A., Popovic,D., Primus,B., Pu,L.-L., Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F., Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J., Sanders,W., Savery,G., Scherer,S., Scott,G., Shatman,S., Shen,H., Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D., Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J., Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Umani,K., Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F., Williams,G., Willson,R., Wleczek,R., Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,D., Yoon,V., Yu,F., Zhang,J., Zhou,X., Zhou,X., Zhao,S., Dunn,D., von Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O., Weinstock,G. and Gibbs,R.A.

Direct Submission
Unpublished
2 (bases 1 to 235381)
Worley,K.C.
Direct Submission
Submitted (26-MAY-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 235381)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (15-NOV-2002) Human Genome Sequencing Center, Department

of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On Nov 15, 2002 this sequence version replaced gi:23265039. The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information

Center project name: GYVO
Center clone name: CH230-154H6
----- Summary Statistics

Assembly program: Phrap; version 0.990329
Consensus quality: 220971 bases at least Q40
Consensus quality: 224061 bases at least Q30
Consensus quality: 226217 bases at least Q20
Estimated insert size: 230107; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
* as soon as it is available and the accession number will
* be preserved.

* 1 97368: contig of 97368 bp in length
* 97369 97468: gap of unknown length
* 97469 233531: contig of 136063 bp in length
* 233532 233631: gap of unknown length
* 233632 235381: contig of 1750 bp in length.

FEATURES

source
1. .235381
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clones="CH230-154H6"
1. .2102
/note="wgs end extension
clone_end:Sp6"
4409. .5070
/note="clone_boundary
clone_end:Sp6
site:
end sequence:BH362740"
97369. 97468
/estimated_length=unknown
129461. .129613
/note="clone_boundary
clone_end:T7
site:
end sequence:BH362739"
230457. .231609
/note="wgs_end extension
clone_end:T7"
231949. .233531
/note="wgs_end extension
clone_end:T7"

233532. .233631
/estimated_length=unknown

ORIGIN

Query Match 87.2%; Score 21.8; DB 14; Length 235381;
Best Local Similarity 92.0%; Pred. No. 12;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAGAGCGCAGACTGCTCTTCC 25
|||||
45019 AAAAAGAGCGTGAATGCTCTTCC 45043
|||||

RESULT 47
AC103495/c

LOCUS
DEFINITION
AC103495
AC103495.7 GI:30520399
HTG: HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
Rattus norvegicus (Norway rat)
SOURCE
ORGANISM
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Murinae; Rattus.
1 (bases 1 to 237321)
Muzny, D.Marie., Metzker, M.Lee., Abramson, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,
Chacko, V., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
Egan, A., Escoto, M., Evans, C., Evans, C.A., Falls, T., Fan, G.,
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, I., Garza, M.,
Gebregiorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, M.,
Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K.,
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hoques, M.,
Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
Karpach, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
Lorenshewa, L., Loulsegged, H., Lozado, R.J., Lu, X., Ma, J.,
Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A.,
Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
Mawhiney, S., McLeod, M.P., McNeill, T.Z., Meenen, E.,
Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,
Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
Nwaokeme, O., Okwuonu, G., Olarnpusagoon, A., Pal, S., Parks, K.,
Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,
Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L.,
Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R.,
Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J.,
Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,
Shetty, J., Shvartabeyn, A., Sisson, I., Sitter, C.D., Smajls, D.,
Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,
Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C.,
Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,
Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J.,
Wang, O., Wang, S., Warren, R., Warren, R., Wei, X., White, F.,
Williams, G., Willson, R., Wlezyk, R., Wooden, H., Worley, K.,
Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,

Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O.,
Weinstock, G. and Gibbs, R. A.

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

Submitted (25-NOV-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 237321)

Rat Genome Sequencing Consortium.

TITLE

JOURNAL

Submitted (10-MAY-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

On May 10, 2003 this sequence version replaced gi:24635679.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GHTM

Center clone name: CH230-31E1

----- Summary Statistics

Assembly program: Atlas 3.0;

Consensus quality: 220013 bases at least Q40

Consensus quality: 222844 bases at least Q30

Consensus quality: 224875 bases at least Q20

Estimated insert size: 231039; sum-of-contigs estimation

Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).

* NOTE: This is a 'working draft' sequence. It currently

* consists of 4 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

* as soon as it is available and the accession number will

* be preserved.

* 1 232262: contig of 232262 bp in length

* 232263 232362: gap of unknown length

* 232363 234000: contig of 1638 bp in length

* 234001 234100: gap of unknown length

* 234101 235489: contig of 1389 bp in length

* 235490 235589: gap of unknown length

* 235590 237321: contig of 1732 bp in length.

Location/Qualifiers

1. .237321

/organism="Rattus norvegicus"

/mol_type="genomic DNA"

/db_xref="taxon:10116"

/clone="CH230-31E1"

1. .1834

/notes="wgs end extension

clone end:Sp6"

4516..5301

misc_feature

misc_feature

/note="clone boundary
clone end:Sp6
site:EcoRI
end_sequence:BH320543"
complement(124058..231882)
/note="clone boundary
clone end:T7
site:EcoRI
end_sequence:BH320542"
129379..131737
/note="wgs end extension
clone end:T7"
232263..232362
/estimated_length=unknown
234001..234100
/estimated_length=unknown
235490..235589
/estimated_length=unknown

ORIGIN

Query Match 87.2%; Score 21.8; DB 14; Length 237321;
Best Local Similarity 92.0%; Pred. NO. 12;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAAGAGCGCAGACTGGTCTTCC 25

|||||
Db 115298 AAAAAAGAGCGCTGACTGCTTCC 115274

RESULT 48

AC126306

AC126306 238355 bp DNA linear HTG 09-OCT-2002
Rattus norvegicus clone CH230-207N13, *** SEQUENCING IN PROGRESS
***, 6 unordered pieces.

ACCESSION

AC126306.2 GI:23196210

VERSION HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.

KEYWORDS Rattus norvegicus (Norway rat)

SOURCE Rattus norvegicus

ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

Sciurognathi; Muridea; Muridae; Murinae; Rattus.

1 (bases 1 to 238355)

REFERENCE

AUTHORS

Murphy, M., Morris, K., Morris, S., Munida, M., Murphy, M., Nair, L.,

Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,

Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,

Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,

Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,

Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,

Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,

Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, J.,

Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,

Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,

Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,

Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,

Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G.,

Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,

Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,

Gebrgeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, M.,

Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,

Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,

Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hognes, M.,

Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,

Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jollivet, A.,

Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,

Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,

Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,

Lorensuhewa, L., Loulasegh, H., Lozano, R. J., Lu, X., Ma, J.,

Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A.,

Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,

Mawhinney, S., McLeod, M. P., McNeill, T. Z., Meenen, E.,

Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,

Morgan, M., Morris, K., Morris, S., Munida, M., Murphy, M., Nair, L.,

Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
 Nwokeneme, O., Okwuonu, G., Olarnunsgoon, A., Pal, S., Parks, K.,
 Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkoch, C.,
 Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L.,
 Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R.,
 Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
 Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J.,
 Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,
 Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajs, D.,
 Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,
 Steimle, M., Strong, R., Sutton, A., Svatek, A., Taber, F., Taylor, C.,
 Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Umani, K.,
 Valas, R., Vera, V., Villaseana, D., Waldron, L., Walker, B., Wang, J.,
 Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, P.,
 Williams, G., Willson, R., Wleczek, R., Wooden, H., Worley, K.,
 Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
 Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
 Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,
 Weinstein, G. and Gibbs, R.A.

TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL

Unpublished
 2 (bases 1 to 238355)
 Worley, K.C.
 Direct Submission
 Submitted (05-JUL-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA

REFERENCE
 AUTHORS
 TITLE
 JOURNAL

3 (bases 1 to 238355)
 Rat Genome Sequencing Consortium.
 Direct Submission
 Submitted (09-OCT-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Sep 19, 2002 this sequence version replaced gi:21699124.
 The sequence in this assembly is a combination of BAC based reads
 and whole genome shotgun sequencing reads assembled using Atlas
 (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). As a result, the
 sequence may extend beyond the ends of the clone and there may be
 contigs that consist entirely of whole genome shotgun sequence
 reads. Both end sequences and whole genome shotgun sequence only
 contigs will be indicated in the feature table.

COMMENT

----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: GZIV
 Center clone name: CH230-207N13
 ----- Summary Statistics
 Assembly program: Phrap; version 0.990329
 Consensus quality: 185501 bases at least Q40
 Consensus quality: 188832 bases at least Q30
 Consensus quality: 190590 bases at least Q20
 Estimated insert size: 233967; sum-of-contigs estimation
 Quality coverage: 3x in Q20 bases; sum-of-contigs estimation

 * NOTE: Estimated insert size may differ from sequence length
 (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 6 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 * 1 211754: contig of 211754 bp in length
 * 211755 211854: gap of unknown length
 * 211855 228024: contig of 16170 bp in length
 * 228025 228124: gap of unknown length
 * 228125 229182: contig of 1058 bp in length
 * 229183 229282: gap of unknown length

* 229283 230367: contig of 1085 bp in length
 * 230368 230467: gap of unknown length
 * 230468 232860: contig of 2393 bp in length
 * 232861 232960: gap of unknown length
 * 232961 238355: contig of 5395 bp in length.
 FEATURES
 source
 1..238355
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-207N13"
 misc_feature
 1..1097
 /note="wgs_contig"
 misc_feature
 26398..29670
 /note="wgs_contig"
 gap
 211755..211854
 /estimated_length=unknown
 misc_feature
 211855..213065
 /note="wgs_contig"
 gap
 228025..228124
 /estimated_length=unknown
 gap
 229183..229282
 /estimated_length=unknown
 gap
 230368..230467
 /estimated_length=unknown
 gap
 232861..232960
 /estimated_length=unknown
 ORIGIN
 Query Match 87.2%; Score 21.8; DB 14; Length 238355;
 Best Local Similarity 92.0%; Pred. No. 12;
 Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 AAAAAAGCGCAGCTGCTCTCC 25
 Db 226727 AAAAAAGCGCTGCTCTCTCC 226751
 RESULT 49
 AC128432/c
 LOCUS
 DEFINITION
 Rattus norvegicus clone CH230-103P9, *** SEQUENCING IN PROGRESS
 ***, 13 unordered pieces.
 AC128432
 AC128432.3 GI:25007946
 VERSION
 HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
 KEYWORDS
 Rattus norvegicus (Norway rat)
 SOURCE
 Rattus norvegicus
 ORGANISM
 Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muridea; Muridae; Murinae; Rattus.
 1 (bases 1 to 239591)
 Muzny, D., Marie, Metzker, M., Lee, A., Adams, C., Alder, J.,
 Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
 Anyalebechi, V., Ayodeji, M., Baca, E., Baden, H.,
 Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
 Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
 Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
 Cardenas, V., Carter, K., Cavazos, I., Caesar, H., Center, A.,
 Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
 Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
 Davila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
 Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
 Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
 Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G.,
 Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
 Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
 Gebregorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
 Gunaratne, P., Haaland, W., Hamil, C., Hamilton, N., Hernandez, J.,
 Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
 Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M.,
 Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,
 Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,

ORIGIN

Query Match 87.2%; Score 21.8; DB 14; Length 239591;
 Best Local Similarity 92.0%; Pred. No. 12;
 Matches 23; Conservatives 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAAGAGCGAGACTGCTCTTCC 25

Db 198883 AAAAAAGAGCGTGTCTCTTCC 198859

RESULT 50

AC103505/c

LOCUS AC103505 239946 bp DNA linear HTG 10-MAY-2003
 DEFINITION Rattus norvegicus clone CH230-94F6, *** SEQUENCING IN PROGRESS ***,
 2 unordered pieces.

ACCESSION AC103505.6 GI:30520584

VERSION HTG: HTGS PHASE1; HTGS DRAFT; HTGS_ENRICHED.

KEYWORDS Rattus norvegicus (Norway rat)

SOURCE

ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muridea; Muridae; Murinae; Rattus.

REFERENCE

AUTHORS Muzny, D. Marie., Metzker, M. Lee., Abramson, S., Adams, C., Alder, J.,

Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
 Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
 Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
 Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
 Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
 Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,
 Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
 Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
 Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
 Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
 Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Faves, K.,
 Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G.,
 Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
 Fraser, C. M., Gabis, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
 Gregeorge, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
 Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K.,
 Harvey, Y., Havlak, P., Haves, A., Henderson, N., Hernandez, J.,
 Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M.,
 Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,
 Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
 Karpach, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
 Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
 Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
 Lorensueta, L., Loulseg, H., Lozado, R. J., Lu, X., Ma, J.,
 Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A.,
 Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
 Mawhney, S., McLeod, M. P., McNeill, T. Z., Meenen, E.,
 Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
 Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,
 Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
 Nwaokeme, O., Okwuonu, G., Olarnpungoon, A., Pal, S., Parks, K.,
 Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkoch, C.,
 Plopper, F., Poindestex, A., Popovic, D., Primus, E., Pu, L., L.,
 Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R.,
 Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
 Rives, C., Rodkey, T., Rojars, A., Rose, M., Rose, R., Ruiz, S. J.,
 Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,
 Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smajda, D.,
 Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,
 Stealmie, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C.,
 Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,
 Valas, R., Vera, V., Villaseana, D., Waldron, L., Walker, B., Wang, J.,
 Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
 Williams, G., Willson, R., Wleczky, R., Woodson, H., Worley, K.,
 Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
 Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
 Niederhausen, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O.,

Weinstock, G. and Gibbs, R. A.

Direct Submission

Unpublished

2 (bases 1 to 239946)

Worley, K. C.

Direct Submission

Submitted (25-NOV-2001) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 239946)

Rat Genome Sequencing Consortium.

Direct Submission

Submitted (10-MAY-2003) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA

On May 10, 2003 this sequence version replaced gi:25139811.

The sequence in this assembly is a combination of BAC based reads
 and whole genome shotgun sequencing reads assembled using Atlas
 (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
 in the feature table below represents a scaffold in the Atlas
 assembly (a 'contig-scaffold'). Within each contig-scaffold,
 individual sequence contigs are ordered and oriented, and separated
 by sized gaps filled with Ns to the estimated size. The sequence
 may extend beyond the ends of the clone and there may be sequence
 contigs within a contig-scaffold that consist entirely of whole
 genome shotgun sequence reads. Both end sequences and whole genome
 shotgun sequence only contigs will be indicated in the feature
 table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: http://www.hgsc.bcm.tmc.edu/

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GHTY

Center clone name: CH230-94F6

----- Summary Statistics

Assembly program: Atlas 3.0;

Consensus quality: 221977 bases at least Q40

Consensus quality: 224834 bases at least Q30

Consensus quality: 226977 bases at least Q20

Estimated insert size: 238332; sum-of-contigs estimation

Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length

* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).

* NOTE: This is a 'working draft' sequence. It currently

* consists of 2 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

* as soon as it is available and the accession number will

* be preserved.

* 1 237949: contig of 237949 bp in length

* 237950 238049: gap of unknown length

* 238050 239946: contig of 1897 bp in length.

FEATURES

source

1. 239946
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-94F6"

misc_feature

3340..4123
 /note="clone_boundary
 clone_end:Sp6
 site:EcoRI"

misc_feature

end_sequence:BH330542"
 complement(235281..236203)
 /note="clone_boundary
 clone_end:T7
 site:EcoRI
 end_sequence:BH330539"

```
misc_feature      complement(235421..236183)
                  /note="clone_boundary
clone_end:T7
site:ECORI
end_sequence:BH330539"
gap              237950..238049
                  /estimated_length=unknown
ORIGIN
Query Match      87.2%; Score 21.8; DB 14; Length 239946;
Best Local Similarity 92.0%; Pred. No. 12;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 AAAAAAGAGCGGAGACTGCTCTTCC 25
    |||||
Db 41607 AAAAAAGAGCGTGAAGCTCTTCC 41583
    |||||

Search completed: February 3, 2006, 21:27:55
Job time : 892 secs
```

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 3, 2006, 10:53:33 ; Search time 203.111 Seconds
(without alignments)
820.326 Million cell updates/sec

Title: US-10-719-900-35

Perfect score: 25

Sequence: 1 aaaaagagcgagactggtcttcc 25

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database :

N Geneseq_21.*

1: Geneseqn1980s.*

2: Geneseqn1980s.*

3: Geneseqn2000s.*

4: Geneseqn2001as.*

5: Geneseqn2001bs.*

6: Geneseqn2002as.*

7: Geneseqn2002bs.*

8: Geneseqn2003as.*

9: Geneseqn2003bs.*

10: Geneseqn2003cs.*

11: Geneseqn2003ds.*

12: Geneseqn2004as.*

13: Geneseqn2004bs.*

14: Geneseqn2005as.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	21.8	87.2	36599	13	Abd33506 Murine ca
2	20.2	80.8	1650	10	Ad448813 Human gen
3	19.8	79.2	68842	14	Adz13146 Murine ca
4	19.4	77.6	30752	11	Acn44344 Mouse gen
5	19.4	77.6	90091	10	Adc85509 Mouse mCG
6	19.4	77.6	90100	9	Ada03029 Mouse mCG
7	19.4	77.6	90100	10	Adb72767 Mouse mCG
8	19.4	77.6	90100	12	Adm74624 Murine ca
9	19.2	76.8	1276	8	Ad55112 Murine lc
10	19.2	76.8	1276	8	Acc45096 Murine lc
11	19.2	76.8	1392	8	Acc45096 Murine lc
12	19.2	76.8	29956	9	Adc85509 Mouse mCG
13	19.2	76.8	29956	10	Adc85509 Mouse mCG
14	19.2	76.8	29956	10	Adc85509 Mouse mCG
15	19.2	76.8	29956	12	Adm74558 Murine ca
16	19.2	76.8	110000	6	Adv34995 Murine cd
17	19.2	76.8	123192	13	Adv34995 Murine cd
18	18.8	75.2	43445	12	Adh18345 Human apo
19	18.8	75.2	43445	12	Ado32886 Human apo


```
XX 14-AUG-2002; 2002WO-US025765.
XX
XX 14-AUG-2001; 2001US-0312147P.
XX
XX 01-NOV-2001; 2001US-0346382P.
XX
XX 26-NOV-2001; 2001US-0333347P.
XX
XX (GEHO ) GEN HOSPITAL CORP.
XX
XX (FARB ) BAYER AG.
XX
XX Woelf C, D'urso D, Befort K, Costigan M;
XX
XX WPI; 2003-268312/26.
XX
XX GENBANK; NM_016337.
XX
XX New composition comprising two or more isolated polypeptides, useful for
XX preparing a medicament for treating pain in an animal.
XX
XX Example 1; Page; 1017pp; English.
XX
XX The invention discloses a composition comprising two or more isolated rat
XX or human polynucleotides or a polynucleotide which represents a fragment,
XX derivative or allelic variation of the nucleic acid sequence. Also
XX claimed are a vector comprising the novel polynucleotide, a host cell
XX comprising the vector, a method for identifying a nucleotide sequence
XX which is differentially regulated in an animal subjected to pain and a
XX kit to perform the method, an array, a method for identifying an agent
XX that increases or decreases the expression of the polynucleotide sequence
XX that is differentially expressed in neuronal tissue of a first animal
XX subjected to pain, a method for identifying a compound which regulates
XX the expression of a polynucleotide sequence which is differentially
XX expressed in an animal subjected to pain, a method for identifying a
XX compound that regulates the activity of one or more of the
XX polynucleotides, a method for producing a pharmaceutical composition, a
XX method for identifying a compound or small molecule that regulates the
XX activity in an animal of one or more of the polypeptides given in the
XX specification, a method for identifying a compound useful in treating
XX pain and a pharmaceutical composition comprising the one or more
XX polypeptides or their antibodies. The polynucleotide or the compound that
XX modulates its activity is useful for preparing a medicament for treating
XX pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
XX injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
XX therapy). The sequence presented is a human DNA (described in Table 3 of
XX the specification) which encodes one of the polypeptides of the invention
XX which is differentially expressed during pain. Note: The sequence data
XX for this patent did not form part of the printed specification, but was
XX obtained in electronic form directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences.
XX
XX SQ Sequence 1650 BP; 418 A; 488 C; 443 G; 301 T; 0 U; 0 Other;
XX
XX Query Match 80.8%; Score 20.2; DB 10; Length 1650;
XX Best Local Similarity 88.0%; Pred. No. 15;
XX Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
XX
XX QY 1 AAAAAAGCGCAGACTGGTCTTCC 25
XX ||||| ||||| ||||| ||||| |||||
XX Db 1393 AAAAAAGCGCAGACTGGTCTTCC 1369
XX
XX RESULT 3
XX ADZ13146
XX ID ADZ13146 standard; DNA; 68842 BP.
XX
XX AC ADZ13146;
XX
XX XX 16-JUN-2005 (first entry)
XX
XX DE Murine cancer-associated genomic DNA #57.
XX
XX XX Diagnosis; DNA microarray; microarray; biochip; cancer; neoplasm;
XX cytostatic; gene; db.
XX
XX XX
```

```
OS Mus sp.
XX WO2005031001-A2.
XX
XX PD 07-APR-2005.
XX
XX PF 23-SEP-2004; 2004WO-US031617.
XX
XX PR 23-SEP-2003; 2003US-00669920.
XX
XX PA (CHIR ) CHIRON CORP.
XX
XX PI Morris DW, Malandro MS;
XX
XX DR WPI; 2005-273395/28.
XX
XX PT Nucleic acid array useful for detecting cancer associated nucleic acid,
XX comprises two or more nucleic acid probes.
XX
XX PS Disclosure; SEQ ID NO 666; 198pp; English.
XX
XX CC The invention relates to a nucleic acid array for detecting a cancer
XX associated (CA) nucleic acid, comprising two or more nucleic acid probes.
XX The invention also relates to a peptide array comprising two or more
XX isolated polypeptides encoded by a CA nucleic acid sequence, a compound
XX that binds to a polypeptide, an isolated antibody or its fragment which
XX binds to a polypeptide, which is prepared by immunizing a host animal
XX with a composition comprising the polypeptide or its antigen binding
XX fragment and collecting cells from the host expressing antibodies against
XX the antigen or its antigen binding fragment, a composition comprising the
XX antibody and a carrier, a method of screening for anticancer activity, a
XX method of detecting a CA nucleic acid, a method of inhibiting expression of a CA
XX method of treating cancer and a method of inhibiting expression of a CA
XX nucleic acid in a cell. The CA nucleic acids are useful for detecting CA
XX nucleic acids. The antibody is useful for detecting the presence or
XX absence of cancer cells in an individual which involves contacting cells
XX from the individual with the antibody and detecting a complex of a CA
XX protein from the cancer cells and the antibody, where the detection of
XX the complex correlates with the presence of cancer cells in the
XX individual. The composition is useful for inhibiting growth of cancer
XX cells in an individual or for delivering a therapeutic agent to cancer
XX cells in an individual. The invention is also useful for diagnosing
XX cancer, for treating cancer and for inhibiting expression of a CA gene in
XX a cell. This sequence represents murine cancer-associated genomic DNA of
XX the invention.
XX
XX SQ Sequence 68842 BP; 18666 A; 15495 C; 16367 G; 18314 T; 0 U; 0 Other;
XX
XX Query Match 79.2%; Score 19.8; DB 14; Length 68842;
XX Best Local Similarity 91.3%; Pred. No. 49;
XX Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
XX QY 3 AAAAAGCGCAGACTGGTCTTCC 25
XX ||||| ||||| ||||| ||||| |||||
XX Db 21274 ATAAAGCGCAGACTGGTCTTCC 21296
XX
XX RESULT 4
XX ACN44344
XX ID ACN44344 standard; DNA; 30752 BP.
XX
XX AC ACN44344;
XX
XX XX 18-NOV-2004 (first entry)
XX
XX DT Mouse genomic sequence MCG17984.
XX
XX DE Cytostatic; carcinoma; lymphoma; cancer; murine; gene; ss.
XX
XX KW Mus musculus.
XX
XX OS WO2003073826-A2.
XX
XX PN
```

```

PD 12-SEP-2003.
XX
XX
PF 28-FEB-2003; 2003WO-US0062235.
XX
XX
PR 01-MAR-2002; 2002US-00087192.
XX
XX
PA (SAGR-) SAGRES DISCOVERY.
XX
XX
PI Morris DW;
XX
XX
DR WPI; 2003-328604/31.
XX
XX
PT Recombinant nucleic acid useful for diagnosis and treatment of carcinoma
PT comprises a nucleotide sequence.
XX
XX
PS Claim 1; SEQ ID NO 745; Opp; English.
XX
XX
CC The present invention relates to novel DNA and protein sequences which
CC are associated with carcinomas. The sequences are useful for: (i) for
CC screening drug candidates; (ii) for screening of bioactive agent capable
CC of binding to Carcinoma Associated Protein (CAP); (iii) for screening of
CC a bioactive agent capable of modulating the activity of CAP; (iv) for
CC evaluating the effect of a candidate carcinoma drug; (v) for diagnosing
CC carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating
CC carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip;
CC (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for
CC determining Carcinoma Associated (CA) gene copy number. In addition, the
CC CA genes are useful as DNA vaccines and the CAP are useful as markers of
CC carcinoma including lymphoma. The present sequence is one such CA coding
CC sequence. Note: This patent is an equivalent to basic patent
CC US200218236A1, for which no sequence data was published
XX
XX
SQ Sequence 30752 BP; 8068 A; 6394 C; 7002 G; 7660 T; 0 U; 1628 Other;

Query Match 77.6%; Score 19.4; DB 11; Length 30752;
Best Local Similarity 95.2%; Pred. No. 66;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 AAGAGCGCAGACTGCTCTTCC 25
Db 3215 AAGAGCGCAGACTGCTCTTCC 3235

RESULT 5
ADC85509
ID ADC85509 standard; DNA; 90091 BP.
XX
XX
AC ADC85509;
XX
XX
DT 01-JAN-2004 (first entry)
XX
XX
DE Mouse mCG10516 genomic sequence.
XX
XX
KW Cytostatic; gene therapy; vaccine; cancer; carcinoma-associated gene; CA;
KW secreted; transmembrane; intracellular; ds.
XX
XX
OS Mus sp.
XX
XX
PN WO2003045230-A2.
XX
XX
PD 05-JUN-2003.
XX
XX
PF 02-DEC-2002; 2002WO-US039582.
XX
XX
PR 30-NOV-2001; 2001US-000997722.
XX
XX
PA (SAGR-) SAGRES DISCOVERY.
XX
XX
PI Morris DW, Engelhard EK;
XX
XX
DR WPI; 2003-513603/48.
XX
XX
PT New recombinant nucleic acid comprising a nucleotide sequence of any of

the carcinoma-associated (CA) genes, useful for screening for drug
candidates for diagnosing or treating carcinomas.
XX
XX
PS Claim 1; SEQ ID NO 295; 983pp; English.
XX
XX
CC The invention relates to a recombinant nucleic acid comprising a
CC nucleotide sequence selected from any of the fully defined carcinoma-
CC associated (CA) genes from the 50 tables given in the specification. The
CC CA proteins are secreted, transmembrane or intracellular proteins. The
CC recombinant nucleic acids are useful for screening for drug candidates
CC for diagnosing or treating carcinomas. Sequences given in ADC85215-
CC ADC85514 represent CA genes of the invention.
XX
XX
SQ Sequence 90091 BP; 22131 A; 13818 C; 15427 G; 24918 T; 0 U; 13797 Other;

Query Match 77.6%; Score 19.4; DB 10; Length 90091;
Best Local Similarity 95.2%; Pred. No. 81;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 AAGAGCGCAGACTGCTCTTCC 25
Db 50171 AAGAGCGCAGACTGCTCTTCC 50191

RESULT 6
ADA03029
ID ADA03029 standard; DNA; 90100 BP.
XX
XX
AC ADA03029;
XX
XX
DT 06-NOV-2003 (first entry)
XX
XX
DE Mouse mCG10516 carcinoma associated gene, SEQ ID NO:1547.
XX
XX
KW Mouse; murine; carcinoma associated; oncogene; carcinoma; cancer; breast;
KW prostate; lymphoma; leukaemia; cytostatic; gene therapy; drug screening;
XX
XX
OS Mus sp.
XX
XX
PN WO2003057146-A2.
XX
XX
PD 17-JUL-2003.
XX
XX
PF 26-DEC-2002; 2002WO-US041414.
XX
XX
PR 26-DEC-2001; 2001US-00035832.
XX
XX
PA (SAGR-) SAGRES DISCOVERY.
XX
XX
PI Morris DW;
XX
XX
DR WPI; 2003-587068/55.
XX
XX
PT New recombinant nucleic acid encoding carcinoma associated protein,
PT useful for preparing compositions for treating carcinomas.
XX
XX
PS Claim 1; SEQ ID NO 1547; 245pp; English.
XX
XX
CC The invention relates to recombinant carcinoma associated (CA) nucleic
CC acid sequences from mouse and human (ADA01482-ADA03094), and to
CC recombinant carcinoma associated proteins (CAP) encoded by them. The
CC invention also encompasses expression vectors and host cells comprising a
CC CA nucleic acid, a polypeptide (especially an antibody) that specifically
CC binds to the protein, and a biochip comprising CA nucleic acid or
CC fragments thereof. The sequences of the invention were identified using
CC oncogenic retroviruses, which insert into the genome of the host organism
CC at random. Many of these do not carry transduced host oncogenes or
CC pathogenic trans-acting viral genes, meaning that cancer incidence is a
CC direct consequence of the effects of proviral integration into host
CC protooncogenes. The CA nucleic acid sequences can be used to diagnose
CC carcinoma (especially breast cancer, prostate cancer, lymphoma or
CC leukaemia) or a propensity to carcinoma by determination of the sequence

```


CC of a CA gene, or by determination of CA gene expression in particular
 CC tissues. CA nucleic acids, proteins and antibodies are also useful as
 CC therapeutic agents and in screening and evaluating drug candidates. The
 CC present sequence represents a specifically claimed murine CA nucleic acid
 CC sequence of the invention. Note: The complete sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 90100 BP; 22131 A; 13818 C; 15428 G; 24917 T; 0 U; 13806 Other;

Query Match 77.6%; Score 19.4; DB 9; Length 90100;
 Best Local Similarity 95.2%; Pred. No. 81;
 Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 AAGAGCGCAGACTGCTCTTCC 25
 |||||
 DB 50180 AAGAGCGCAGACTGCTCTTCC 50200

RESULT 7
 ADB72767
 ID ADB72767 standard; DNA; 90100 BP.

XX AC ADB72767;
 XX DT 04-DEC-2003 (first entry)

XX Mouse MCG10516 gene.

XX mouse; ds; cytostatic; gene therapy; vaccine; carcinoma; lymphomas;
 KW cancer; neoplasm; adenocarcinoma, sarcoma; gene.

XX Mus sp.

XX WO2003008583-A2.

XX 30-JAN-2003.

XX 26-DEC-2001; 2001WO-US051291.

XX 02-MAR-2001; 2001US-00798586.

XX 23-OCT-2001; 2001US-00004113.

XX 08-NOV-2001; 2001US-00052482.

XX 30-NOV-2001; 2001US-00997722.

XX 20-DEC-2001; 2001US-00034650.

XX (SAGR-) SAGRES DISCOVERY.

XX Morris DW, Engelhard EK;

XX WPI; 2003-239337/23.

XX New recombinant nucleic acid, useful for treating carcinomas, lymphomas,
 XX cancers, neoplasm, adenocarcinoma, or sarcomas.

XX Claim 1; SEQ ID NO 595; 2304pp; English.
 XX The invention relates to a novel recombinant nucleic acid comprising a
 XX nucleotide sequence selected from any of the 660 sequences fully defined
 XX in the specification. A polynucleotide of the invention has cytostatic
 XX activity, and may have a use in gene therapy, or in a vaccine. The
 XX recombinant nucleic acids and polypeptides are useful for treating
 XX carcinomas, e.g. lymphomas, cancers, neoplasm, adenocarcinoma, and
 XX sarcomas. The present sequence represents a mouse gene of the invention.

XX SQ Sequence 90100 BP; 22131 A; 13818 C; 15428 G; 24917 T; 0 U; 13806 Other;

Query Match 77.6%; Score 19.4; DB 10; Length 90100;
 Best Local Similarity 95.2%; Pred. No. 81;
 Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 AAGAGCGCAGACTGCTCTTCC 25

DB 50180 AAGAGCGCAGACTGCTCTTCC 50200

RESULT 8
 ADM74624

ID ADM74624 standard; DNA; 90100 BP.

XX ADM74624;

XX 01-JUL-2004 (first entry)

XX Murine carcinoma associated (CA) nucleic acid #148.

XX Mouse; carcinoma associated nucleic acid; CA nucleic acid; gene; ds;
 KW carcinoma associated protein; CAP; carcinoma; leukaemia; lymphoma;
 KW cytostatic.

XX Mus musculus.

XX US2004072154-A1.

XX 15-APR-2004.

XX 30-NOV-2001; 2001US-00997722.

XX 22-DEC-2000; 2000US-00747377.

XX 02-MAR-2001; 2001US-00798586.

XX (MORRIS) MORRIS D W.

XX (ENGELHARD) ENGELHARD E K.

XX Morris DW, Engelhard EK;

XX WPI; 2004-328562/30.

XX New carcinoma associated gene or protein, useful for preparing a
 XX composition for diagnosing or treating carcinoma e.g., leukemia or
 XX lymphoma.

XX Claim 1; SEQ ID NO 295; 29pp; English.

XX The invention relates to new recombinant nucleic acids. The invention
 CC also relates to a host cell comprising a recombinant nucleic acid or
 CC expression vector, an expression vector comprising a recombinant nucleic
 CC acid, a recombinant protein, a method of screening for drug candidates, a
 CC method of screening for a bioactive agent capable of binding to a
 CC carcinoma associated protein (CAP) encoded by a nucleotide sequence, a
 CC method of screening for a bioactive agent capable of modulating the
 CC activity of a CAP, a method of evaluating the effect of a candidate
 CC carcinoma drug, a method of diagnosing carcinoma, a method for inhibiting
 CC the activity of a CAP, a method of treating carcinomas, a method of
 CC neutralising the effect of a CAP and a method of diagnosing carcinoma or
 CC propensity to carcinoma. A method of evaluating the effect of a candidate
 CC carcinoma drug comprises administering the drug to a patient, removing a
 CC cell sample from the patient and determining alterations in the
 CC expression or activation of a gene comprising the nucleotide sequence. A
 CC method of diagnosing carcinoma comprises determining the expression of
 CC one or more genes comprising the nucleic acid sequence in a first tissue
 CC type of a first individual and comparing the expression of the gene from
 CC a second normal tissue from the first individual or a second
 CC unaffected individual, where a difference in the expression indicates
 CC that the first individual has carcinoma. A method of inhibiting the
 CC activity of a CAP comprises binding an inhibitor to the CAP. Treating
 CC carcinomas comprises administering to a patient an inhibitor of CAP.
 CC Neutralising the effect of a CAP comprises contacting an agent specific
 CC for the CAP. The polypeptide specifically binds to the protein encoded by
 CC the nucleic acid. It comprises an antibody that specifically binds to the
 CC protein encoded by the nucleic acid. The nucleic acids are useful for
 CC preparing a composition for diagnosing or treating carcinoma e.g.,
 CC leukaemia or lymphoma. This sequence represents a murine carcinoma
 CC associated (CA) nucleic acid of the invention. Note: The sequence data
 CC for this patent did not form part of the printed specification but was

```
CC obtained in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 90100 BP; 22131 A; 13818 C; 15428 G; 24917 T; 0 U; 13806 Other;

Query Match 77.6%; Score 19.4; DB 12; Length 90100;
Best Local Similarity 95.2%; Pred. No. 81;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 AAGAGCGCAGACTGCTCTTCC 25
Db 50180 AAGAGCGCAGACTGCTCTTCC 50200

RESULT 9
AAD5112/c
ID AAD51112 standard; DNA; 1276 BP.
XX
AC AAD51112;
XX
XX 07-AUG-2003 (first entry)
XX
DE Murine lck promoter DNA fragment.
XX
KW Transgenic; biotechnology; agriculture; murine; promoter; ds.
XX
OS Mus sp.
XX
PN WO2003022040-A2.
XX
PD 20-MAR-2003.
XX
PF 12-SEP-2002; 2002WO-US029130.
XX
PR 13-SEP-2001; 2001US-0322031P.
PR 09-JAN-2002; 2002US-0347782P.
XX
XX (CALY ) CALIFORNIA INST OF TECHNOLOGY.
XX
PI Baltimore D, Hong EU, Lois-Caballe C, Pease S;
XX
XX WPI; 2003-300976/29.
XX
XX Producing a transgenic animal for commercial use, comprises transfecting
XX a packaging cell line with retroviral construct, recovering recombinant
XX retrovirus from the cell line and infecting embryonic cell with the
XX recombinant virus.
XX
XX Example 1; Fig 19B; 76pp; English.
XX
XX The invention relates to a method for producing a transgenic animal for
XX commercial use, which comprises transfecting a packaging cell line with
XX retroviral construct, recovering recombinant retrovirus from the cell
XX line and infecting embryonic cell with the recombinant virus. The method
XX is useful in producing transgenic animals using retroviral constructs
XX engineered to carry a transgene of interest. The transgenic animals may
XX find use in commercial applications like biotechnology and agriculture.
XX The present sequence is murine lck promoter DNA fragment used to
XX illustrate the method of the invention
XX
SQ Sequence 1276 BP; 302 A; 259 C; 381 G; 334 T; 0 U; 0 Other;

Query Match 76.8%; Score 19.2; DB 8; Length 1276;
Best Local Similarity 87.5%; Pred. No. 45;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAAGAGCGCAGACTGCTCTTCC 25
Db 66 AAAAAGAGCGCAGACTGCTCTTCC 43

RESULT 10
AAC45096/c
ID AAC45096 standard; DNA; 1276 BP.
XX
AC AAC45096;
XX
XX 10-JUN-2003 (first entry)
XX
XX Murine lck promoter partial nucleotide sequence SEQ ID NO:6.
XX
XX Transgenic animal; transgenic bird; transgenic fish; transgene;
XX retroviral construct; lentiviral; long terminal repeat; LTR;
XX biotechnology; agriculture; murine; lck promoter; ds.
XX
XX Mus sp.
XX
XX WO2003022228-A2.
XX
XX 20-MAR-2003.
XX
XX 12-SEP-2002; 2002WO-US029157.
XX
XX 13-SEP-2001; 2001US-0322031P.
XX 09-JAN-2002; 2002US-0347782P.
XX
XX (CALY ) CALIFORNIA INST OF TECHNOLOGY.
XX
XX Baltimore D, Hong EU, Lois-Caballe C, Pease S;
XX
XX WPI; 2003-301005/29.
XX
XX Producing a transgenic bird or fish for commercial use, comprises
XX transfecting a packaging cell line with retroviral construct, recovering
XX recombinant retrovirus from the cell and infecting bird or fish egg with
XX the recombinant virus.
XX
XX Example 1; Fig 19; 68pp; English.
XX
XX The present invention describes a method for producing a transgenic bird
XX or fish. The method comprises transfecting a packaging cell line with a
XX retroviral construct, recovering recombinant retroviral particles from
XX the packaging cell line, and infecting a bird or a fish egg with the
XX recombinant retroviral particles. The retroviral construct comprises the
XX R and U5 sequences from a 5' lentiviral long terminal repeat (LTR) and a
XX self-inactivating 3' lentiviral LTR. Also described is a transgenic bird
XX or fish made by the above method and whose genome comprises a proviral
XX DNA that has a self-inactivating 3' lentiviral LTR. The method is useful
XX in producing transgenic animals, particularly transgenic birds and fish,
XX using retroviral constructs engineered to carry a transgene of interest.
XX The method is used to introduce the gene of choice into animals in order
XX to confer upon them desired attributes. The transgenic animals may find
XX use in commercial applications like biotechnology and agriculture. The
XX present sequence represents a murine lck promoter partial nucleotide
XX sequence, which is used in an example from the present invention
XX
SQ Sequence 1276 BP; 302 A; 259 C; 381 G; 334 T; 0 U; 0 Other;

Query Match 76.8%; Score 19.2; DB 8; Length 1276;
Best Local Similarity 87.5%; Pred. No. 45;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAAGAGCGCAGACTGCTCTTCC 25
Db 66 AAAAAGAGCGCAGACTGCTCTTCC 43

RESULT 11
ACA36353/c
ID ACA36353 standard; DNA; 1392 BP.
XX
AC ACA36353;
XX
XX 19-JUN-2003 (first entry)
XX
XX Prokaryotic essential gene #18010.
```

```
XX Antisense; ds; prokaryotic essential gene; cell proliferation;
KW drug design; gene.
XX
XX Listeria monocytogenes.
XX WO200277183-A2.
XX
XX 03-OCT-2002.
XX
XX 21-MAR-2002; 2002WO-US0009107.
XX
XX 21-MAR-2001; 2001US-00815242.
XX
XX 06-SEP-2001; 2001US-00948993.
XX
XX 25-OCT-2001; 2001US-0342923P.
XX
XX 08-FEB-2002; 2002US-00072851.
XX
XX 06-MAR-2002; 2002US-0362699P.
XX
XX (ELIT-) ELITRA PHARM INC.
XX
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
XX Wall D, Trawick JD, Carr GU, Yamamoto R, Forsyth RA, Xu HH;
XX
XX WPI; 2003-029926/02.
XX
XX P-PSDB; ABU32483.
XX
XX New antisense nucleic acids, useful for identifying proteins or screening
XX PT for homologous nucleic acids required for cellular proliferation to
XX PT isolate candidate molecules for rational drug discovery programs.
XX
XX Claim 14; SEQ ID NO 24223; 1766pp; English.
XX
XX The invention relates to an isolated nucleic acid comprising any one of
XX CC the 6213 antisense sequences given in the specification where expression
XX CC of the nucleic acid inhibits proliferation of a cell. Also included are:
XX CC (1) a vector comprising a promoter operably linked to the nucleic acid
XX CC encoding a polypeptide whose expression is inhibited by the antisense
XX CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX CC polypeptide or its fragment whose expression is inhibited by the
XX CC antisense nucleic acid; (4) an antibody capable of specifically binding
XX CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
XX CC proliferation or the activity of a gene in an operon required for
XX CC proliferation; (7) identifying a compound that influences the activity of
XX CC the gene product or that has an activity against a biological pathway
XX CC required for proliferation, or that inhibits cellular proliferation; (8)
XX CC identifying a gene required for cellular proliferation or the biological
XX CC pathway in which a proliferation-required gene or its gene product lies
XX CC or a gene on which the test compound that inhibits proliferation of an
XX CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
XX CC compound's activity; (11) a culture comprising strains in which the gene
XX CC product is overexpressed or underexpressed; (12) determining the extent
XX CC to which each of the strains is present in a culture or collection of
XX CC strains; or (13) identifying the target of a compound that inhibits the
XX CC proliferation of an organism. The antisense nucleic acids are useful for
XX CC identifying proteins or screening for homologous nucleic acids required
XX CC for cellular proliferation to isolate candidate molecules for rational
XX CC drug discovery programs, or for screening homologous nucleic acids
XX CC required for proliferation in cells other than S. aureus, S. typhimurium,
XX CC K. pneumoniae or P. aeruginosa. The present sequence is one of the target
XX CC prokaryotic essential genes. Note: The sequence data for this patent did
XX CC not form part of the printed specification, but was obtained in
XX CC electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 1392 BP; 442 A; 230 C; 346 G; 374 T; 0 U; 0 Other;
XX
XX Query Match 76.8%; Score 19.2; DB 8; Length 1392;
XX Best Local Similarity 87.5%; Pred. No. 46;
XX Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
XX
XX 1 AAAAAAGAGCGCAGACTGGTCTTC 24
XX |||||||||||||||||||||
XX 132 AAAAAAGAGCGCAGACTGATTTTC 109
XX
XX Db
```

```
RESULT 12
ADA02963/c
ID ADA02963 standard; DNA; 29956 BP.
XX
XX AC ADA02963;
XX
XX 06-NOV-2003 (first entry)
XX
XX Mouse Lck carcinoma associated gene, SEQ ID NO:1481.
XX
XX Mouse; murine; carcinoma associated; oncogene; carcinoma; cancer; breast;
XX KW prostate; lymphoma; leukaemia; cytostatic; gene therapy; drug screening;
XX KW gene; ds.
XX
XX OS Mus sp.
XX
XX WO2003057146-A2.
XX
XX 17-JUL-2003.
XX
XX 26-DEC-2002; 2002WO-US041414.
XX
XX 26-DEC-2001; 2001US-00035832.
XX
XX (SAGR-) SAGRES DISCOVERY.
XX
XX Morris DW;
XX
XX WPI; 2003-587068/55.
XX
XX New recombinant nucleic acid encoding carcinoma associated protein,
XX PT useful for preparing compositions for treating carcinomas.
XX
XX Claim 1; SEQ ID NO 1481; 245pp; English.
XX
XX The invention relates to recombinant carcinoma associated (CA) nucleic
XX CC acid sequences from mouse and human (ADA01482-ADA03094), and to
XX CC recombinant carcinoma associated proteins (CAP) encoded by them. The
XX CC invention also encompasses expression vectors and host cells comprising a
XX CC CA nucleic acid, a polypeptide (especially an antibody) that specifically
XX CC binds to the protein, and a biochip comprising CA nucleic acid or
XX CC fragments thereof. The sequences of the invention were identified using
XX CC oncogenic retroviruses, which insert into the genome of the host organism
XX CC at random. Many of these do not carry transduced host oncogenes or
XX CC pathogenic trans-acting viral genes, meaning that cancer incidence is a
XX CC direct consequence of the effects of proviral integration into host
XX CC protooncogenes. The CA nucleic acid sequences can be used to diagnose
XX CC carcinoma (especially breast cancer, prostate cancer, lymphoma or
XX CC leukaemia) or a propensity to carcinoma by determination of the sequence
XX CC of a CA gene, or by determination of CA gene expression in particular
XX CC tissues. CA nucleic acids, proteins and antibodies are also useful as
XX CC therapeutic agents and in screening and evaluating drug candidates. The
XX CC present sequence represents a specifically claimed murine CA nucleic acid
XX CC sequence of the invention. Note: The complete sequence data for this
XX CC patent did not form part of the printed specification, but was obtained
XX CC in electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 29956 BP; 6634 A; 6708 C; 6677 G; 7953 T; 0 U; 1984 Other;
XX
XX Query Match 76.8%; Score 19.2; DB 9; Length 29956;
XX Best Local Similarity 87.5%; Pred. No. 82;
XX Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
XX
XX 2 AAAAAAGAGCGCAGACTGGTCTTC 25
XX |||||||||||||||||||||
XX Db 8863 AAAAAAGAGCGCAGCTGCTCTTC 8840
XX
XX RESULT 13
XX ADB72701/c
```

```

ID ADB72701 standard; DNA; 29956 BP.
XX
AC ADB72701;
XX
DT 04-DEC-2003 (first entry)
XX
DE Mouse Lck gene.
XX
KW mouse; ds; cytotstatic; gene therapy; vaccine; carcinoma; lymphomas;
KW cancer; neoplasm; adenocarcinoma; sarcoma; gene.
XX
OS Mus sp.
XX
PN WO2003008583-A2.
XX
PD 30-JAN-2003.
XX
PF 26-DEC-2001; 2001WO-US051291.
XX
PR 02-MAR-2001; 2001US-00798586.
PR 23-OCT-2001; 2001US-00004113.
PR 08-NOV-2001; 2001US-00052482.
PR 30-NOV-2001; 2001US-00997722.
PR 20-DEC-2001; 2001US-00034650.
XX
PA (SAGR-) SAGRES DISCOVERY.
XX
PI Morris DW, Engelhard EK;
XX
PT New recombinant nucleic acid, useful for treating carcinomas, lymphomas,
PT cancers, neoplasm, adenocarcinoma, or sarcomas.
XX
PS Claim 1; SEQ ID NO 529; 2304pp; English.
XX
CC The invention relates to a novel recombinant nucleic acid comprising a
CC nucleotide sequence selected from any of the 660 sequences fully defined
CC in the specification. A polynucleotide of the invention has cytostatic
CC activity, and may have a use in gene therapy, or in a vaccine. The
CC recombinant nucleic acids and polypeptides are useful for treating
CC carcinomas, e.g. lymphomas, cancers, neoplasm, adenocarcinoma, and
CC sarcomas. The present sequence represents a mouse gene of the invention.
XX
SQ Sequence 29956 BP; 6634 A; 6708 C; 6677 G; 7953 T; 0 U; 1984 Other;

Query Match 76.8%; Score 19.2; DB 10; Length 29956;
Best Local Similarity 87.5%; Pred. No. 82;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAAGAGCGAGCTGCTCTTCC 25
DB 8863 AAAAAGAGCACTGACTGCTCTTCC 8840

RESULT 14
ADC85443/C
ID ADC85443 standard; DNA; 29956 BP.
XX
AC ADC85443;
XX
DT 01-JAN-2004 (first entry)
XX
DE Mouse Lck genomic sequence.
XX
KW Cytostatic; gene therapy; vaccine; cancer; carcinoma-associated gene; CA;
KW secreted; transmembrane; intracellular; ds.
XX
OS Mus sp.
XX
PN WO2003045230-A2.
XX
PD 05-JUN-2003.

ADB72701 standard; DNA; 29956 BP.
XX
AC ADB72701;
XX
DT 04-DEC-2003 (first entry)
XX
DE Mouse Lck gene.
XX
KW mouse; ds; cytotstatic; gene therapy; vaccine; carcinoma; lymphomas;
KW cancer; neoplasm; adenocarcinoma; sarcoma; gene.
XX
OS Mus sp.
XX
PN WO2003008583-A2.
XX
PD 30-JAN-2003.
XX
PF 26-DEC-2001; 2001WO-US051291.
XX
PR 02-MAR-2001; 2001US-00798586.
PR 23-OCT-2001; 2001US-00004113.
PR 08-NOV-2001; 2001US-00052482.
PR 30-NOV-2001; 2001US-00997722.
PR 20-DEC-2001; 2001US-00034650.
XX
PA (SAGR-) SAGRES DISCOVERY.
XX
PI Morris DW, Engelhard EK;
XX
PT New recombinant nucleic acid, useful for treating carcinomas, lymphomas,
PT cancers, neoplasm, adenocarcinoma, or sarcomas.
XX
PS Claim 1; SEQ ID NO 529; 2304pp; English.
XX
CC The invention relates to a novel recombinant nucleic acid comprising a
CC nucleotide sequence selected from any of the 660 sequences fully defined
CC in the specification. A polynucleotide of the invention has cytostatic
CC activity, and may have a use in gene therapy, or in a vaccine. The
CC recombinant nucleic acids and polypeptides are useful for treating
CC carcinomas, e.g. lymphomas, cancers, neoplasm, adenocarcinoma, and
CC sarcomas. The present sequence represents a mouse gene of the invention.
XX
SQ Sequence 29956 BP; 6634 A; 6708 C; 6677 G; 7953 T; 0 U; 1984 Other;

Query Match 76.8%; Score 19.2; DB 10; Length 29956;
Best Local Similarity 87.5%; Pred. No. 82;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAAGAGCGAGCTGCTCTTCC 25
DB 8863 AAAAAGAGCACTGACTGCTCTTCC 8840

RESULT 15
ADM74558/C
ID ADM74558 standard; DNA; 29956 BP.
XX
AC ADM74558;
XX
DT 01-JUL-2004 (first entry)
XX
DE Murine carcinoma associated (CA) nucleic acid #115.
XX
KW Mouse; carcinoma associated nucleic acid; CA nucleic acid; gene; ds;
KW carcinoma associated protein; CAP; carcinoma; leukaemia; lymphoma;
KW cytostatic.
XX
OS Mus musculus.
XX
PN US2004072154-A1.
XX
PD 15-APR-2004.
XX
PF 30-NOV-2001; 2001US-00997722.
XX
PR 22-DEC-2000; 2000US-00747377.
PR 02-MAR-2001; 2001US-00798586.
XX
PA (MORR/) MORRIS D W.
PA (ENGE/) ENGELHARD E K.
XX
PI Morris DW, Engelhard EK;
XX
PWPI; 2004-328562/30.
XX
PT New carcinoma associated gene or protein, useful for preparing a
PT composition for diagnosing or treating carcinoma e.g., leukemia or
PT lymphoma.
XX
PS Claim 1; SEQ ID NO 229; 29pp; English.
```

Fragment Name	Begin	End
ABA03041_00	1	110000
WP		
ABA03041_01	100001	210000
WP		
ABA03041_02	200001	310000
WP		
ABA03041_03	300001	410000
WP		
ABA03041_04	400001	510000
WP		
ABA03041_05	500001	610000
WP		
ABA03041_06	600001	710000
WP		
ABA03041_07	700001	810000
WP		
ABA03041_08	800001	910000
WP		
ABA03041_09	900001	1010000
WP		
ABA03041_10	1000001	1110000
WP		
ABA03041_11	1100001	1210000
WP		
ABA03041_12	1200001	1310000
WP		
ABA03041_13	1300001	1410000
WP		
ABA03041_14	1400001	1510000
WP		
ABA03041_15	1500001	1610000
WP		
ABA03041_16	1600001	1710000
WP		
ABA03041_17	1700001	1810000
WP		
ABA03041_18	1800001	1910000
WP		
ABA03041_19	1900001	2010000
WP		
ABA03041_20	2000001	2110000
WP		
ABA03041_21	2100001	2210000
WP		

CC of this invention exhibit various activities including neuroleptic,
CC nootropic, antimanic and antidepressant. Furthermore, the screening
CC method used in MPHTS will be automated, such that a large number of test
CC compounds may be rapidly screened with a minimal amount of labour and
CC effort. This polynucleotide is the cDNA sequence of a gene that is
CC differentially expressed in mice in the presence of the therapeutic
CC compound valproate, given in an exemplification of the invention.

XX
SQ Sequence 123192 BP; 37038 A; 25773 C; 25891 G; 34490 T; 0 U; 0 Other;

Query Match 76.8%; Score 19.2; DB 13; Length 123192;

Best Local Similarity 87.5%; Pred. No. 1.1e+02;

Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAAGAGCGAGACTGCTTTC 24

Db 73891 AAAAAAGAGCACTGACTGCTTTC 73914

RESULT 18

ADH18345

ID ADH18345 standard; DNA; 43445 BP.

AC ADH18345;

XX 11-MAR-2004 (first entry)

DE Human apolipoprotein B (ApoB) DNA 3 - SEQ ID 334.

XX apolipoprotein B; ApoB; antiarteriosclerotic; cardiant; antidiabetic;
KW anorectic; lipid; cholesterol metabolism; atherosclerosis;
KW diabetes Type 2; obesity; hyperlipidaemia; cardiovascular; gene therapy;
KW human; ds.

XX Homo sapiens.

OS WO2003097662-A1.

PN 27-NOV-2003.

XX 15-MAY-2003; 2003WO-US015493.

XX 15-MAY-2002; 2002US-00147196.

PR 13-NOV-2002; 2002US-0426234P.

XX (ISIS-) ISIS PHARM INC.

PA Crooke RM, Graham MJ;

XX WPI; 2004-022840/02.

XX New antisense compound, useful for preparing a composition for treating
PT abnormal lipid or cholesterol metabolism, atherosclerosis, diabetes Type
PT 2, obesity, hyperlipidemia or cardiovascular disease.
XX Example 33; SEQ ID NO 334; 405pp; English.

XX The invention relates to a novel antisense compound targeted to a nucleic
CC acid molecule encoding human apolipoprotein B (ApoB) which specifically
CC hybridises with and inhibits the expression of human apolipoprotein B.
CC The compound of the invention demonstrates antiarteriosclerotic,
CC cardiant, antidiabetic and anorectic activities and may be useful for
CC preparing a composition for treating abnormal lipid or cholesterol
CC metabolism, atherosclerosis, diabetes Type 2, obesity, hyperlipidaemia or
CC cardiovascular disease. Furthermore, the compound has gene therapy
CC applications. The current sequence is that of the human ApoB DNA of the
CC invention.

XX Sequence 43445 BP; 12661 A; 9313 C; 9323 G; 12148 T; 0 U; 0 Other;

Query Match 75.2%; Score 18.8; DB 12; Length 43445;

Best Local Similarity 90.9%; Pred. No. 1.4e+02;

Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAAGAGCGAGACTGCTTCT 22
||||| ||||| ||||| |||||
Db 25768 AAAAAAGAGCGTAGCTGCTTCT 25789

RESULT 19

ADO32886

ID ADO32886 standard; DNA; 43445 BP.

XX ADO32886;

XX 12-AUG-2004 (first entry)

XX Human apolipoprotein B (ApoB) gDNA - SEQ ID 334.

XX apolipoprotein B; ApoB; cardiovascular; antiarteriosclerotic;
KW antilipaeamic; antidiabetic; anorectic; cardiant; vasotropic; hypotensive;
KW anabolic; eating disorder; cytostatic; endocrine; vasotropic;
KW neuroprotective; nootropic; lipid; cholesterol metabolism;
KW hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;
KW Von Gierke's disease; lipodystrophy; Cushing's syndrome;
KW sexual ateliotic dwarfism; hyperthyroidism; hypertension;
KW anorexia nervosa; Werner's syndrome; hepatoma; multiple myeloma; uraemia;
KW impotence; obstructive liver disease; Alzheimer's; dementia; diabetes;
KW obesity; atherosclerosis; human; ds; chromosome 2p23-2p24.

XX Homo sapiens.

OS WO2004044181-A2.

PN 27-MAY-2004.

XX 13-NOV-2003; 2003WO-US036411.

XX 13-NOV-2002; 2002US-0426234P.

PR 15-MAY-2003; 2003WO-US015493.

XX (ISIS-) ISIS PHARM INC.

XX Crooke R, Graham M, Lemonidis-Tarbet K, Dobie KW;

XX WPI; 2004-420321/39.

XX Antisense oligonucleotide compound that inhibits expression of mRNA
PT encoding human apolipoprotein B, useful for treating hyperlipidemia,
PT diabetes, obesity, von Gierke's disease, lipodystrophies, Cushing's
PT syndrome.
XX Example 33; SEQ ID NO 334; 483pp; English.

XX The invention relates to a novel antisense compound where the compound
CC hybridises to and inhibits expression of mRNA encoding human
CC apolipoprotein B (ApoB) after 16-24 hours by at least 30% in 80%
CC confluent HepG2 cells in culture at a concentration of 150 nM. The
CC compound of the invention demonstrates cardiovascular,
CC antiarteriosclerotic, antilipaeamic, antidiabetic, anorectic, cardiant,
CC vasotropic, hypotensive, anabolic, eating disorder-related, cytostatic,
CC endocrine, vasotropic, neuroprotective and nootropic activities and may
CC be useful for inhibiting the expression of apolipoprotein B in cells or
CC tissues in vivo in order to address a condition associated with abnormal
CC lipid or cholesterol metabolism. The compound may be useful for
CC decreasing circulating lipoprotein levels, triglyceride levels,
CC cholesterol levels, lipid levels, fatty acid levels, acute phase
CC reactants and chylomicrons and thus may be utilised during treatment of
CC hyperlipoproteinaemia, hyperlipidaemia, hypercholesterolaemia,
CC cardiovascular disorders, Von Gierke's disease, lipodystrophy, Cushing's
CC syndrome, sexual ateliotic dwarfism, hyperthyroidism, hypertension,
CC anorexia nervosa, Werner's syndrome, hepatoma, multiple myeloma, uraemia,
CC impotence, obstructive liver disease, Alzheimer's disease, dementia,
CC diabetes, obesity and atherosclerosis. The current sequence is that of
CC the human apolipoprotein B (ApoB) gDNA - SEQ ID 334 of the invention
CC which is located at chromosome 2p23-2p24.

```
XX SQ Sequence 43445 BP; 12661 A; 9314 C; 9322 G; 12148 T; 0 U; 0 Other;
Query Match 75.2%; Score 18.8; DB 12; Length 43445;
Best Local Similarity 90.9%; Pred. No. 1.4e+02;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAAGCGCAGACTGGTCT 22
||||| ||||| ||||| ||||| |||||
Db 25768 AAAAAAGCGCAGACTGGTCT 25789

RESULT 20
ACN44704
ID ACN44704 standard; DNA; 39405 BP.
XX AC ACN44704;
XX
DT 18-NOV-2004 (first entry)
XX
DE Mouse genomic sequence MCG2509.
XX
KW Cytostatic; carcinoma; lymphoma; cancer; murine; gene; ss.
XX
OS Mus musculus.
XX
PN WO2003073826-A2.
XX
PD 12-SEP-2003.
XX
PF 28-FEB-2003; 2003WO-US006235.
XX
PR 01-MAR-2002; 2002US-00087192.
XX
PA (SAGR-) SAGRES DISCOVERY.
XX
PI Morris DW;
XX
DR WPI; 2003-328604/31.
XX
PT Recombinant nucleic acid useful for diagnosis and treatment of carcinoma
PT comprises a nucleotide sequence.
XX
PS Claim 1; SEQ ID NO 1285; Opp; English.
XX
CC The present invention relates to novel DNA and protein sequences which
CC are associated with carcinomas. The sequences are useful for: (i) for
CC screening drug candidates; (ii) for screening of bioactive agent capable
CC of binding to Carcinoma Associated Protein (CAP); (iii) for screening of
CC a bioactive agent capable of modulating the activity of CAP; (iv) for
CC evaluating the effect of a candidate carcinoma drug; (v) for diagnosing
CC carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating
CC carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip;
CC (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for
CC determining Carcinoma Associated (CA) gene copy number. In addition, the
CC CA genes are useful as DNA vaccines and the CAP are useful as markers of
CC carcinoma including lymphoma. The present sequence is one such CA coding
CC sequence. Note: This patent is an equivalent to basic patent
CC US2002182586A1, for which no sequence data was published
XX
SQ Sequence 39405 BP; 10420 A; 8422 C; 8690 G; 11580 T; 0 U; 293 Other;
Query Match 74.4%; Score 18.6; DB 11; Length 39405;
Best Local Similarity 84.0%; Pred. No. 1.7e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAAGCGCAGACTGGTCTTC 25
||||| ||||| ||||| ||||| |||||
Db 9488 AAAAAAGCGCAGACTGGTCTTC 9512

RESULT 21
ACN44796
```

```
ID ACN44796 standard; DNA; 39790 BP.
XX AC ACN44796;
XX
DT 18-NOV-2004 (first entry)
XX
DE Mouse genomic sequence MCG4192.
XX
KW Cytostatic; carcinoma; lymphoma; cancer; murine; gene; ss.
XX
OS Mus musculus.
XX
PN WO2003073826-A2.
XX
PD 12-SEP-2003.
XX
PF 28-FEB-2003; 2003WO-US006235.
XX
PR 01-MAR-2002; 2002US-00087192.
XX
PA (SAGR-) SAGRES DISCOVERY.
XX
PI Morris DW;
XX
DR WPI; 2003-328604/31.
XX
PT Recombinant nucleic acid useful for diagnosis and treatment of carcinoma
PT comprises a nucleotide sequence.
XX
PS Claim 1; SEQ ID NO 1423; Opp; English.
XX
CC The present invention relates to novel DNA and protein sequences which
CC are associated with carcinomas. The sequences are useful for: (i) for
CC screening drug candidates; (ii) for screening of bioactive agent capable
CC of binding to Carcinoma Associated Protein (CAP); (iii) for screening of
CC a bioactive agent capable of modulating the activity of CAP; (iv) for
CC evaluating the effect of a candidate carcinoma drug; (v) for diagnosing
CC carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating
CC carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip;
CC (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for
CC determining Carcinoma Associated (CA) gene copy number. In addition, the
CC CA genes are useful as DNA vaccines and the CAP are useful as markers of
CC carcinoma including lymphoma. The present sequence is one such CA coding
CC sequence. Note: This patent is an equivalent to basic patent
CC US2002182586A1, for which no sequence data was published
XX
SQ Sequence 39790 BP; 9126 A; 9461 C; 9319 G; 8925 T; 0 U; 2959 Other;
Query Match 74.4%; Score 18.6; DB 11; Length 39790;
Best Local Similarity 84.0%; Pred. No. 1.7e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAAGCGCAGACTGGTCTTC 25
||||| ||||| ||||| ||||| |||||
Db 15595 AAAAAAGCGCAGACTGGTCTTC 15619

RESULT 22
ADA02627/c
ID ADA02627 standard; DNA; 47115 BP.
XX AC ADA02627;
XX
DT 06-NOV-2003 (first entry)
XX
DE Mouse Flt3 carcinoma associated gene, SEQ ID NO:1145.
XX
KW Mouse; murine; carcinoma associated; oncogene; carcinoma; cancer; breast;
KW prostate; lymphoma; leukaemia; cytostatic; gene therapy; drug screening;
KW gene; ds.
XX
OS Mus sp.
```

PN WO2003057146-A2.
XX
PD 17-JUL-2003.
XX
PF 26-DEC-2002; 2002WO-US041414.
XX
PR 26-DEC-2001; 2001US-00035832.
XX
PA (SAGR-) SAGRES DISCOVERY.
XX
PI Morris DW;
XX
XX WPI; 2003-587068/55.
XX
XX New recombinant nucleic acid encoding carcinoma associated protein,
PT useful for preparing compositions for treating carcinomas.
XX
XX Claim 1; SEQ ID NO 1145; 245pp; English.
XX
XX The invention relates to recombinant carcinoma associated (CA) nucleic
CC acid sequences from mouse and human (ADA01482-ADA03094), and to
CC recombinant carcinoma associated proteins (CAP) encoded by them. The
CC invention also encompasses expression vectors and host cells comprising a
CC CA nucleic acid, a polypeptide (especially an antibody) that specifically
CC binds to the protein, and a biochip comprising CA nucleic acid or
CC fragments thereof. The sequences of the invention were identified using
CC oncogenic retroviruses, which insert into the genome of the host organism
CC at random. Many of these do not carry transduced host oncogenes or
CC pathogenic trans-acting viral genes, meaning that cancer incidence is a
CC direct consequence of the effects of proviral integration into host
CC protooncogenes. The CA nucleic acid sequences can be used to diagnose
CC carcinoma (especially breast cancer, prostate cancer, lymphoma or
CC leukaemia) or a propensity to carcinoma by determination of the sequence
CC of a CA gene, or by determination of CA gene expression in particular
CC tissues. CA nucleic acids, proteins and antibodies are also useful as
CC therapeutic agents and in screening and evaluating drug candidates. The
CC present sequence represents a specifically claimed murine CA nucleic acid
CC sequence of the invention. Note: The complete sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 47115 BP; 11483 A; 10181 C; 10540 G; 13012 T; 0 U; 1899 Other;

Query Match 74.4%; Score 18.6; DB 9; Length 47115;
Best Local Similarity 84.0%; Pred. No. 1.8e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAAGAGCGAGCTGCTCTTCC 25
Db ||||||||| ||||||||| ||||||||| ||||||||| |||||||||
5679 AAAAAAGAGCGAGCTGCTCTTCC 5655

RESULT 23
ADB72365/c
ID ADB72365 standard; DNA; 47115 BP.
XX
AC ADB72365;
XX
DT 04-DEC-2003 (first entry)
XX
XX Mouse Flt3 gene.
DE
XX mouse; ds; cytostatic; gene therapy; vaccine; carcinoma; lymphomas;
KW cancer; neoplasm; adenocarcinoma; sarcoma; gene.
XX
OS Mus sp.
XX
XX WO2003008583-A2.
PN
XX 30-JAN-2003.
PD
XX 26-DEC-2001; 2001WO-US051291.
PF

XX 02-MAR-2001; 2001US-00798586.
PR 23-OCT-2001; 2001US-00004113.
PR 08-NOV-2001; 2001US-00052482.
PR 30-NOV-2001; 2001US-00997722.
PR 20-DEC-2001; 2001US-00034650.
XX
PA (SAGR-) SAGRES DISCOVERY.
XX
XX Morris DW, Engelhard EK;
PI WPI; 2003-239337/23.
XX
XX New recombinant nucleic acid, useful for treating carcinomas, lymphomas,
PT cancers, neoplasm, adenocarcinoma, or sarcomas.
XX
XX Claim 1; SEQ ID NO 193; 2304pp; English.
XX
XX The invention relates to a novel recombinant nucleic acid comprising a
CC nucleotide sequence selected from any of the 660 sequences fully defined
CC in the specification. A polynucleotide of the invention has cytostatic
CC activity, and may have a use in gene therapy, or in a vaccine. The
CC recombinant nucleic acids and polypeptides are useful for treating
CC carcinomas, e.g. lymphomas, cancers, neoplasm, adenocarcinoma, and
CC sarcomas. The present sequence represents a mouse gene of the invention.
XX
SQ Sequence 47115 BP; 11483 A; 10181 C; 10540 G; 13012 T; 0 U; 1899 Other;

Query Match 74.4%; Score 18.6; DB 10; Length 47115;
Best Local Similarity 84.0%; Pred. No. 1.8e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAAGAGCGAGCTGCTCTTCC 25
Db ||||||||| ||||||||| ||||||||| ||||||||| |||||||||
5679 AAAAAAGAGCGAGCTGCTCTTCC 5655

RESULT 24
ADE95875/c
ID ADE95875 standard; DNA; 47115 BP.
XX
AC ADE95875;
XX
DT 12-FEB-2004 (first entry)
XX
XX Mouse Flt3 gene genomic DNA sequence.
DE
XX cancer diagnosis; cancer treatment; carcinoma; cytostatic; gene therapy;
KW lymphoma; breast cancer; prostate cancer; leukaemia; ds; mouse; murine;
KW Flt3.
XX
OS Mus sp.
XX
XX WO2003039484-A2.
PN
XX 15-MAY-2003.
PD
XX 08-NOV-2002; 2002WO-US036071.
PF
XX 08-NOV-2001; 2001US-00052482.
PR
XX (SAGR-) SAGRES DISCOVERY.
PA
XX Morris DW, Engelhard EK;
PI WPI; 2003-441462/41.
XX
XX New carcinoma associated nucleic acids and proteins, useful for screening
PT drug candidates, or for diagnosing and treating carcinomas, e.g.
PT lymphoma, breast cancer, prostate cancer or leukemia.
XX
XX Claim 1; SEQ ID NO 133; 793pp; English.
XX

CC This invention relates to novel recombinant nucleic acids for use in
CC diagnosis and treatment of cancer, especially carcinomas, as well as the
CC use of compositions in screening methods. The compositions of the
CC invention may have cytostatic activity whilst the disclosed sequences may
CC be useful for gene therapy. The carcinoma associated nucleic acids and
CC proteins are useful for diagnosing and treating carcinomas, for example
CC lymphoma, breast cancer, prostate cancer or leukaemia, or for screening
CC drug candidates or bioactive agents capable of binding to, or modulating
CC the activity of, a carcinoma associated protein. The present sequence is
CC the genomic DNA sequence of the mouse Flt3 gene which is a carcinoma
CC associated gene of the invention.

XX
SQ Sequence 47115 BP; 11483 A; 10181 C; 10539 G; 13012 T; 0 U; 1900 Other;

Query Match 74.4%; Score 18.6; DB 10; Length 47115;
Best Local Similarity 84.0%; Pred. No. 1.8e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAAGAGCGAGACTGGCTCTTCC 25
|||||
Db 5679 AAAAAAGAGCGAGCTGCTCTTCC 5655

RESULT 25
ACN44048/c
ID ACN44048 standard; DNA; 49281 BP.
AC ACN44048;
XX
XX 18-NOV-2004 (first entry)
XX Mouse genomic sequence MCG13312.
DE
XX Cytostatic; carcinoma; lymphoma; cancer; murine; gene; ss.
KW Mus musculus.

OS
XX
XX W02003073826-A2.
XX
XX 12-SEP-2003.
XX
XX 28-FEB-2003; 2003WO-US006235.
XX
XX 01-MAR-2002; 2002US-00087192.
XX (SAGR-) SAGRES DISCOVERY.
PA
XX Morris DW;
XX
XX WPI; 2003-328604/31.
XX Recombinant nucleic acid useful for diagnosis and treatment of carcinoma
PT comprises a nucleotide sequence.

XX Claim 1; SEQ ID NO 301; Opp; English.
XX
XX The present invention relates to novel DNA and protein sequences which
XX are associated with carcinomas. The sequences are useful for: (i) for
XX screening drug candidates; (ii) for screening of bioactive agent capable
XX of binding to Carcinoma Associated Protein (CAP); (iii) for screening of
XX a bioactive agent capable of modulating the activity of CAP; (iv) for
XX evaluating the effect of a candidate carcinoma drug; (v) for diagnosing
XX carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating
XX carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip;
XX (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for
XX determining Carcinoma Associated (CA) gene copy number. In addition, the
XX CA genes are useful as DNA vaccines and the CAP are useful as markers of
XX carcinoma including lymphoma. The present sequence is one such CA coding
XX sequence. Note: This patent is an equivalent to basic patent
XX US2002182586A1, for which no sequence data was published

XX Sequence 49281 BP; 12534 A; 8091 C; 8876 G; 13885 T; 0 U; 5895 Other;

Query Match 74.4%; Score 18.6; DB 11; Length 49281;
Best Local Similarity 84.0%; Pred. No. 1.8e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAAGAGCGAGACTGGCTCTTCC 25
|||||
Db 11432 AAAAAAGAGCGAGCTGCTCTTCC 11408

RESULT 26
ADQ59515/c
ID ADQ59515 standard; DNA; 64482 BP.
XX
XX ADQ59515;
XX
XX 07-OCT-2004 (first entry)
XX Human cancer-associated (CA) gene sequence SEQ ID NO:151.

XX human; cancer-associated gene; cancer-associated protein; cytostatic;
KW gene therapy; vaccine; tyrosine kinase antagonist;
KW G-protein coupled receptor antagonist; cancer; lymphoma; gene; ds.

XX Homo sapiens.
XX W02004058288-A1.
XX
XX 15-JUL-2004.
XX
XX 15-DEC-2003; 2003WO-US040082.
XX
XX 17-DEC-2002; 2002US-00322696.

XX (SAGR-) SAGRES DISCOVERY INC.
XX
XX Morris DW, Malandro MS;
XX WPI; 2004-543349/52.
XX P-PSDB; ADQ59517.

XX New cancer-associated nucleic acid for diagnosing, preventing or treating
PT cancer (e.g. lymphoma) or for screening agents that may be used for
PT treating or preventing cancer.
XX Disclosure; SEQ ID NO 151; 143pp; English.

XX The present invention describes human cancer-associated (CA) nucleotide
CC sequences (1). Also described: (1) an expression vector comprising (1);
CC (2) a host cell comprising (1) or the expression vector; (3) a microarray
CC for detecting a CA nucleic acid; (4) an isolated polypeptide encoded
CC within an open reading frame of a CA sequence; (5) an isolated antibody,
CC or its antigen binding fragment, that binds to the above polypeptide; (6)
CC a hybridoma that produces the monoclonal antibody described above; (7) a
CC pharmaceutical composition comprising the antibody and a pharmaceutical
CC excipient; (8) a kit for detecting or diagnosing cancer cells, comprising
CC the above (monoclonal) antibody or polynucleotide that selectively
CC hybridises to any of the polynucleotide sequences mentioned above; (9)
CC methods for diagnosing cancer or for detecting the presence or absence of
CC cancer cells in an individual; (10) a method for inhibiting growth of
CC cancer cells in an individual; (11) a method for delivering a therapeutic
CC agent to cancer cells in an individual; (12) an electronic library
CC comprising the polynucleotide or polypeptide, or their fragments,
CC mentioned above; (13) a method of screening for anticancer activity; (14)
CC methods for detecting cancer associated with expression of a polypeptide
CC or the presence of the antibody in a test cell or serum sample; (15) a
CC method for screening for a bioactive agent capable of modulating the
CC activity of a CA protein encoded by the above nucleic acid molecule; and
CC (16) a method for treating cancers. (1) has cytostatic activity, and can
CC be used in gene therapy, in vaccines, as a tyrosine kinase antagonist,
CC and as a G-protein coupled receptor antagonist. The compositions and
CC methods of the present invention can be used for diagnosing, preventing
CC and treating cancer, especially lymphomas. They may also be used in
CC screening for agents that may be used for treating or preventing cancer.

CC The present sequence represents a human CA gene sequence, which is given
CC in the exemplification of the present invention. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 64482 BP; 15643 A; 14684 C; 15303 G; 17456 T; 0 U; 1396 Other;

Query Match 74.4%; Score 18.6; DB 12; Length 64482;
Best Local Similarity 84.0%; Pred. No. 1.9e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 AAAAAAGAGCGAGACTGGTCTTCC 25

Db 53861 AATAAAGAGCACTGACTGTTCTTCC 53837

RESULT 27

ADZ13896/c

ID ADZ13896 standard; DNA; 64482 BP.

XX AC ADZ13896;

XX DT 16-JUN-2005 (first entry)

XX DE Murine cancer-associated genomic DNA #123.

XX KW Diagnosis; DNA microarray; microarray; biochip; cancer; neoplasm;
XX KW cytostatic; gene; ds.

XX OS Mus sp.

XX PN WO2005031001-A2.

XX PD 07-APR-2005.

XX PF 23-SEP-2004; 2004WO-US031617.

XX PR 23-SEP-2003; 2003US-00669920.

XX PA (CHIR) CHIRON CORP.

XX PI Morris DW, Malandro MS;

XX DR WPI; 2005-273395/28.

XX PT Nucleic acid array useful for detecting cancer associated nucleic acid,
XX comprises two or more nucleic acid probes.

XX PS Disclosure; SEQ ID NO 1416; 198pp; English.

XX CC The invention relates to a nucleic acid array for detecting a cancer
XX associated (CA) nucleic acid, comprising two or more nucleic acid probes.
XX The invention also relates to a peptide array comprising two or more
XX isolated polypeptides encoded by a CA nucleic acid sequence, a compound
XX that binds to a polypeptide, an isolated antibody or its fragment which
XX binds to a polypeptide, which is prepared by immunizing a host animal
XX with a composition comprising the polypeptide or its antigen binding
XX fragment and collecting cells from the host expressing antibodies against
XX the antigen or its antigen binding fragment, a composition comprising the
XX antibody and a carrier, a method of screening for anticancer activity, a
XX method of detecting a CA nucleic acid, a method of diagnosing cancer, a
XX method of treating cancer and a method of inhibiting expression of a CA
XX nucleic acid in a cell. The CA nucleic acids are useful for detecting CA
XX nucleic acids. The antibody is useful for detecting the presence or
XX absence of cancer cells in an individual which involves contacting cells
XX from the individual with the antibody and detecting a complex of a CA
XX protein from the cancer cells and the antibody, where the detection of
XX the complex correlates with the presence of cancer cells in the
XX individual. The composition is useful for inhibiting growth of cancer
XX cells in an individual or for delivering a therapeutic agent to cancer
XX cells in an individual. The invention is also useful for diagnosing
XX cancer, for treating cancer and for inhibiting expression of a CA gene in

CC a cell. This sequence represents murine cancer-associated genomic DNA of
CC the invention.

XX SQ Sequence 64482 BP; 15643 A; 14684 C; 15303 G; 17456 T; 0 U; 1396 Other;

Query Match 74.4%; Score 18.6; DB 14; Length 64482;

Best Local Similarity 84.0%; Pred. No. 1.9e+02;

Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 AAAAAAGAGCGAGACTGGTCTTCC 25

Db 53861 AATAAAGAGCACTGACTGTTCTTCC 53837

RESULT 28

ABD33614/c

ID ABD33614 standard; DNA; 89213 BP.

XX AC ABD33614;

XX DT 18-NOV-2004 (first entry)

XX DE Murine cancer-associated (CA) gene MD07-125.

XX KW Mouse; cancer-associated protein; CAP; cancer-associated gene; CA; gene;
XX KW ds; cancer; cytostatic.

XX OS Mus musculus.

XX PN WO2004058146-A2.

XX PD 15-JUL-2004.

XX PF 15-DEC-2003; 2003WO-US040081.

XX PR 17-DEC-2002; 2002US-00322281.

XX PA (SAGR-) SAGRES DISCOVERY INC.

XX PI Morris DW, Malandro MS;

XX DR WPI; 2004-499109/47.

XX PT Novel human cancer associated protein encoded within open reading frame
XX of cancer associated gene, useful as targets for diagnosing cancer.

XX PS Disclosure; SEQ ID NO 843; 182pp; English.

XX CC The invention relates to cancer-associated proteins (CAP) and the cancer-
XX associated (CA) nucleic acids encoding them. The invention also relates
XX to a method for treating cancers involving administering to a patient an
XX inhibitor of CAP, and a method of screening for anticancer activity in a
XX potential drug involving providing a cell that expresses a CA gene,
XX contacting a tissue sample derived from a cancer cell with an anticancer
XX drug candidate and monitoring the effect of the anticancer drug candidate
XX on expression of the CA gene. The CAP proteins are useful for detecting
XX cancer associated with expression of a CAP protein in a test cell sample
XX and for screening for a bioactive agent capable of modulating the
XX activity of a CAP protein. The CA nucleic acids are useful for diagnosing
XX cancer, involving determining the expression of a CA nucleic acid in a
XX tissue. This sequence represents a murine CA gene of the invention. Note:
XX The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 89213 BP; 24812 A; 17038 C; 18589 G; 27735 T; 0 U; 1039 Other;

Query Match 74.4%; Score 18.6; DB 13; Length 89213;

Best Local Similarity 84.0%; Pred. No. 2e+02;

Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 AAAAAAGAGCGAGACTGGTCTTCC 25

Db 53861 AATAAAGAGCACTGACTGTTCTTCC 53837

```
Db      81877 AAAAAAGAGTGCTGCTCTTCC 81853

RESULT 29
AD212547/c
ID      AD212547 standard; DNA; 92099 BP.
XX
AC      AD212547;
XX
DT      16-JUN-2005 (first entry)
XX
DE      Murine cancer-associated genomic DNA #8.
XX
KW      Diagnosis; DNA microarray; microarray; biochip; cancer; neoplasm;
KW      cytostatic; gene; ds.
XX
OS      Mus sp.
XX
PN      WO2005031001-A2.
XX
PD      07-APR-2005.
XX
PF      23-SEP-2004; 2004WO-US031617.
XX
PR      23-SEP-2003; 2003US-00669920.
XX
PA      (CHIR ) CHIRON CORP.
XX
PI      Morris DW, Malandro MS;
XX
WPI; 2005-273395/28.
XX
Nucleic acid array useful for detecting cancer associated nucleic acid,
PT      comprises two or more nucleic acid probes.
XX
PS      Disclosure; SEQ ID NO 67; 198pp; English.
XX
CC      The invention relates to a nucleic acid array for detecting a cancer
CC      associated (CA) nucleic acid, comprising two or more nucleic acid probes.
CC      The invention also relates to a peptide array comprising two or more
CC      isolated polypeptides encoded by a CA nucleic acid sequence, a compound
CC      that binds to a polypeptide, an isolated antibody or its fragment which
CC      binds to a polypeptide, which is prepared by immunizing a host animal
CC      with a composition comprising the polypeptide or its antigen binding
CC      fragment and collecting cells from the host expressing antibodies against
CC      the antigen or its antigen binding fragment, a composition comprising the
CC      antibody and a carrier, a method of screening for anticancer activity, a
CC      method of detecting a CA nucleic acid, a method of diagnosing cancer, a
CC      method of treating cancer and a method of inhibiting expression of a CA
CC      nucleic acid in a cell. The CA nucleic acids are useful for detecting CA
CC      nucleic acids. The antibody is useful for detecting the presence or
CC      absence of cancer cells in an individual which involves contacting cells
CC      from the individual with the antibody and detecting a complex of a CA
CC      protein from the cancer cells and the antibody, where the detection of
CC      the complex correlates with the presence of cancer cells in the
CC      individual. The composition is useful for inhibiting growth of cancer
CC      cells in an individual or for delivering a therapeutic agent to cancer
CC      cells in an individual. The invention is also useful for diagnosing
CC      cancer, for treating cancer and for inhibiting expression of a CA gene in
CC      a cell. This sequence represents murine cancer-associated genomic DNA of
CC      the invention.
XX
SQ      Sequence 92099 BP; 23678 A; 20842 C; 20864 G; 25593 T; 0 U; 1122 Other;

Query Match      74.4%; Score 18.6; DB 14; Length 92099;
Best Local Similarity 84.0%; Pred. No. 2e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy      1 AAAAAAGAGCGACGACTGCTCTTCC 25
      |||||
Db      49960 AAAAAAGACATGGACTGCTCTTCC 49936

RESULT 31
ADB72365
ID      ADB72365 standard; DNA; 47115 BP.
XX

RESULT 30
ADA02627
ID      ADA02627 standard; DNA; 47115 BP.
XX
AC      ADA02627;
XX
DT      06-NOV-2003 (first entry)
XX
DE      Mouse Flt3 carcinoma associated gene, SEQ ID NO:1145.
XX
KW      Mouse; murine; carcinoma associated; oncogene; carcinoma; cancer; breast;
KW      prostate; lymphoma; leukaemia; cytostatic; gene therapy; drug screening;
KW      gene; ds.
XX
OS      Mus sp.
XX
PN      WO2003057146-A2.
XX
PD      17-JUL-2003.
XX
PF      26-DEC-2002; 2002WO-US041414.
XX
PR      26-DEC-2001; 2001US-00035832.
XX
PA      (SAGR-) SAGRES DISCOVERY.
XX
PI      Morris DW;
XX
WPI; 2003-597068/55.
XX
New recombinant nucleic acid encoding carcinoma associated protein,
PT      useful for preparing compositions for treating carcinomas.
XX
PS      Claim 1; SEQ ID NO 1145; 245pp; English.
XX
CC      The invention relates to recombinant carcinoma associated (CA) nucleic
CC      acid sequences from mouse and human (ADA01482-ADA03094), and to
CC      recombinant carcinoma associated proteins (CAP) encoded by them. The
CC      invention also encompasses expression vectors and host cells comprising a
CC      CA nucleic acid, a polypeptide (especially an antibody) that specifically
CC      binds to the protein, and a biochip comprising CA nucleic acid or
CC      fragments thereof. The sequences of the invention were identified using
CC      oncogenic retroviruses, which insert into the genome of the host organism
CC      at random. Many of these do not carry transduced host oncogenes or
CC      pathogenic trans-acting viral genes, meaning that cancer incidence is a
CC      direct consequence of the effects of proviral integration into host
CC      protooncogenes. The CA nucleic acid sequences can be used to diagnose
CC      carcinoma (especially breast cancer, prostate cancer, lymphoma or
CC      leukaemia) or a propensity to carcinoma by determination of the sequence
CC      of a CA gene, or by determination of CA gene expression in particular
CC      tissues. CA nucleic acids, proteins and antibodies are also useful as
CC      therapeutic agents and in screening and evaluating drug candidates. The
CC      present sequence represents a specifically claimed murine CA nucleic acid
CC      sequence of the invention. Note: The complete sequence data for this
CC      patent did not form part of the printed specification, but was obtained
CC      in electronic format directly from WIPO at
CC      ftp.wipo.int/pub/published_pct_sequences.
XX
SQ      Sequence 47115 BP; 11483 A; 10181 C; 10540 G; 13012 T; 0 U; 1899 Other;

Query Match      73.6%; Score 18.4; DB 9; Length 47115;
Best Local Similarity 95.0%; Pred. No. 2.2e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      5 AAGAGCGCAGACTGGTCTTCC 24
      |||||
Db      34115 AAGAGCGCAGACTGGTCTTCC 34134

RESULT 31
ADB72365
ID      ADB72365 standard; DNA; 47115 BP.
XX
```

```
AC ADB72365;
XX
DT 04-DEC-2003 (first entry)
XX
XX Mouse Flt3 gene.
XX
XX mouse; ds; cytostatic; gene therapy; vaccine; carcinoma; lymphomas;
KW cancer; neoplasm; adenocarcinoma; sarcoma; gene.
XX
OS Mus sp.
XX
PN WO2003008583-A2.
XX
PD 30-JAN-2003.
XX
PF 26-DEC-2001; 2001WO-US051291.
XX
PR 02-MAR-2001; 2001US-00798586.
PR 23-OCT-2001; 2001US-00004113.
PR 08-NOV-2001; 2001US-00052482.
PR 30-NOV-2001; 2001US-00997722.
PR 20-DEC-2001; 2001US-00034650.
XX
XX (SAGR-) SAGRES DISCOVERY.
XX
XX Morris DW, Engelhard EK;
PI
XX WPI; 2003-239337/23.
XX
XX New recombinant nucleic acid, useful for treating carcinomas, lymphomas,
PT cancers, neoplasm, adenocarcinoma, or sarcomas.
PT
XX Claim 1; SEQ ID NO 193; 2304pp; English.
XX
XX The invention relates to a novel recombinant nucleic acid comprising a
CC nucleotide sequence selected from any of the 660 sequences fully defined
CC in the specification. A polynucleotide of the invention has cytostatic
CC activity, and may have a use in gene therapy, or in a vaccine. The
CC recombinant nucleic acids and polypeptides are useful for treating
CC carcinomas, e.g. lymphomas, cancers, neoplasm, adenocarcinoma, and
CC sarcomas. The present sequence represents a mouse gene of the invention.
XX
XX Sequence 47115 BP; 11483 A; 10181 C; 10540 G; 13012 T; 0 U; 1899 Other;
SQ
Query Match 73.6%; Score 18.4; DB 10; Length 47115;
Best Local Similarity 95.0%; Pred. No. 2.2e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
QY 5 AAGAGCGCAGACTGGCTCTTC 24
Db 34115 AAGAGCGCAGACTGGCTCTTC 34134
RESULT 32
ADE95875
ID ADE95875 standard; DNA; 47115 BP.
XX
AC ADE95875;
XX
DT 12-FEB-2004 (first entry)
XX
DE Mouse Flt3 gene genomic DNA sequence.
XX
XX cancer diagnosis; cancer treatment; carcinoma; cytostatic; gene therapy;
KW lymphoma; breast cancer; prostate cancer; leukaemia; ds; mouse; murine;
KW Flt3.
XX
OS Mus sp.
XX
PN WO2003039484-A2.
XX
PD 15-MAY-2003.
XX
```

```
PF 08-NOV-2002; 2002WO-US036071.
XX
PR 08-NOV-2001; 2001US-00052482.
XX
XX (SAGR-) SAGRES DISCOVERY.
XX
XX Morris DW, Engelhard EK;
XX
XX WPI; 2003-441462/41.
XX
XX New carcinoma associated nucleic acids and proteins, useful for screening
PT drug candidates, or for diagnosing and treating carcinomas, e.g.
PT lymphoma, breast cancer, prostate cancer or leukemia.
XX
XX Claim 1; SEQ ID NO 133; 793pp; English.
XX
XX This invention relates to novel recombinant nucleic acids for use in
CC diagnosis and treatment of cancer, especially carcinomas, as well as the
CC use of compositions in screening methods. The compositions of the
CC invention may have cytostatic activity whilst the disclosed sequences may
CC be useful for gene therapy. The carcinoma associated nucleic acids and
CC proteins are useful for diagnosing and treating carcinomas, for example
CC lymphoma, breast cancer, prostate cancer or leukaemia, or for screening
CC drug candidates or bioactive agents capable of binding to, or modulating
CC the activity of, a carcinoma associated protein. The present sequence is
CC the genomic DNA sequence of the mouse Flt3 gene which is a carcinoma
CC associated gene of the invention.
XX
XX Sequence 47115 BP; 11483 A; 10181 C; 10539 G; 13012 T; 0 U; 1900 Other;
SQ
Query Match 73.6%; Score 18.4; DB 10; Length 47115;
Best Local Similarity 95.0%; Pred. No. 2.2e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
QY 5 AAGAGCGCAGACTGGCTCTTC 24
Db 34115 AAGAGCGCAGACTGGCTCTTC 34134
RESULT 33
ADZ12547
ID ADZ12547 standard; DNA; 92099 BP.
XX
AC ADZ12547;
XX
DT 16-JUN-2005 (first entry)
XX
XX Murine cancer-associated genomic DNA #8.
XX
XX Diagnosis; DNA microarray; microarray; biochip; cancer; neoplasm;
KW cytostatic; gene; ds.
XX
OS Mus sp.
XX
XX WO2005031001-A2.
XX
PD 07-APR-2005.
XX
PF 23-SEP-2004; 2004WO-US031617.
XX
PR 23-SEP-2003; 2003US-00669920.
XX
XX (CHIR ) CHIRON CORP.
XX
XX Morris DW, Malandro MS;
PI
XX WPI; 2005-273395/28.
XX
XX Nucleic acid array useful for detecting cancer associated nucleic acid,
PT comprises two or more nucleic acid probes.
XX
PS Disclosure; SEQ ID NO 67; 198pp; English.
XX
```

CC The invention relates to a nucleic acid array for detecting a cancer
 CC associated (CA) nucleic acid, comprising two or more nucleic acid probes.
 CC The invention also relates to a peptide array comprising two or more
 CC isolated polypeptides encoded by a CA nucleic acid sequence, a compound
 CC that binds to a polypeptide, an isolated antibody or its fragment which
 CC binds to a polypeptide, which is prepared by immunizing a host animal
 CC with a composition comprising the polypeptide or its antigen binding
 CC fragment and collecting cells from the host expressing antibodies against
 CC the antigen or its antigen binding fragment, a composition comprising a
 CC antibody and a carrier, a method of screening for anticancer activity, a
 CC method of detecting a CA nucleic acid, a method of diagnosing cancer, a
 CC method of treating cancer and a method of inhibiting expression of a CA
 CC nucleic acid in a cell. The CA nucleic acids are useful for detecting CA
 CC nucleic acids. The antibody is useful for detecting the presence or
 CC absence of cancer cells in an individual which involves contacting cells
 CC from the individual with the antibody and detecting a complex of a CA
 CC protein from the cancer cells and the antibody, where the detection of
 CC the complex correlates with the presence of cancer cells in the
 CC individual. The composition is useful for inhibiting growth of cancer
 CC cells in an individual or for delivering a therapeutic agent to cancer
 CC cells in an individual. The invention is also useful for diagnosing
 CC cancer, for treating cancer and for inhibiting expression of a CA gene in
 CC a cell. This sequence represents murine cancer-associated genomic DNA of
 CC the invention.

SQ Sequence 92099 BP; 23678 A; 20842 C; 20864 G; 25593 T; 0 U; 1122 Other;
 Query Match 73.6%; Score 18.4; DB 14; Length 92099;
 Best Local Similarity 95.0%; Pred. No. 2.5e+02;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 AAGAGCGCAGACTGGTCTTC 24
 DB 78534 AAGAGCGCAGACTGGTCTTC 78553
 |||||

RESULT 34
 ABZ55378/c
 ID ABZ55378 standard; cDNA; 337 BP.
 AC ABZ55378;
 XX
 XX 28-MAR-2003 (first entry)
 DT
 XX Aspergillus oryzae polynucleotide SEQ ID NO 4491.
 DE
 XX Aspergillus oryzae; fermentation; fungus; industrial; EST;
 KW expressed sequence tag; gene; ss.
 KW
 XX Aspergillus oryzae.
 OS
 XX WO200279476-A1.
 PN
 XX 10-OCT-2002.
 PD
 XX 22-MAR-2002; 2002WO-IB000890.
 PF
 XX 30-MAR-2001; 2001JP-00098371.
 PR
 XX (NAAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.
 PA (NARE-) NAT RES INST BREWING.
 PA (NORQ) NAT FOOD RES INST MIN AGRIC.
 XX
 XX Machida M, Akita O, Kashiwagi Y, Kitamoto K, Horiuchi H;
 PI Takeuchi M, Kobayashi T, Kitamoto N, Gomi K, Abe K;
 PI
 XX WPI; 2003-046817/04.
 DR
 XX Detection of expression of specific Aspergillus genes for monitoring the
 PT fermentation and growth conditions of the fungus, using DNA probes.
 PT
 XX Claim 1; SEQ ID NO 4491; 48pp + Sequence Listing; Japanese.

CC The invention relates to a polynucleotide having any of 6006 specific
 CC sequences (ABZ50888-ABZ56893), which are expressed by a fungus under
 CC specific culture conditions including one or more of eutrophic,
 CC oligotrophic, solid, early germination, alkaline, high temperature, low
 CC temperature or mold culture or polynucleotides stringently hybridising
 CC to these sequences. The polynucleotides are useful for monitoring the
 CC progress of fermentation and the growth conditions of a fungus,
 CC especially of Aspergillus oryzae which is widely used in industrial
 CC fermentation. Also monitoring for fungal contamination. Note: The
 CC sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 337 BP; 76 A; 68 C; 73 G; 120 T; 0 U; 0 Other;
 Query Match 72.8%; Score 18.2; DB 8; Length 337;
 Best Local Similarity 87.0%; Pred. No. 1.1e+02;
 Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAAGAGCGCAGACTGGTCTTC 24
 DB 77 AAAAAGAGCGCTGACAGGTCTTC 55
 |||||

RESULT 35
 ADQ97843/c
 ID ADQ97843 standard; DNA; 92861 BP.
 XX
 AC ADQ97843;
 XX
 XX 07-OCT-2004 (first entry)
 DT
 XX Mouse cancer associated sequence MD11-010, SEQ ID 820.
 DE
 XX Cytostatic; Gene Therapy; cancer; leukemia; lymphoma; Mouse; ds.
 KW
 XX Mus musculus.
 OS
 XX WO2004060304-A2.
 PN
 XX 22-JUL-2004.
 PD
 XX 22-DEC-2003; 2003WO-US041389.
 PF
 XX 27-DEC-2002; 2002US-00330773.
 PR
 XX (SAGR-) SAGRES DISCOVERY INC.
 PA
 XX Morris DW, Malandro MS;
 PI
 XX WPI; 2004-543781/52.
 DR
 XX New isolated cancer associated nucleic acids comprising at least 10
 PT contiguous nucleotides, useful for diagnosing, preventing and/or treating
 PT cancers such as leukemia and lymphoma.
 XX
 PS Claim 1; SEQ ID NO 820; 199pp; English.

CC The present invention relates to cancer associated sequences (ADQ97025-
 CC ADQ98004). The sequences are useful for the diagnosis, prevention and/or
 CC treatment of cancer, such as leukemia and lymphoma. Note: The sequence
 CC data for this patent did not form part of the printed specification, but
 CC was obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 CC
 XX Sequence 92861 BP; 25200 A; 20454 C; 20725 G; 24205 T; 0 U; 2277 Other;

QY 3 AAAAAGAGCGCAGACTGGTCTTC 25
 |||||

Db 1111 AAAAGAGCACTGACTGCTCTTCC 1089

RESULT 36
ADX04494
ID ADX04494 standard; RNA; 95 BP.
XX
AC ADX04494;
XX
DT 05-MAY-2005 (first entry)
XX
DE Rat primary-microRNA (pri-miRNA) mir-333.
XX
XX MicroRNA; ds; RNA interference; Gene silencing; Cytostatic; Antidiabetic;
KW Anorectic; Antilipemic; Antiarteriosclerotic; Hypotensive;
KW Neuroprotective; Nootropic; Antiangiogenic; Anabolic;
KW Eating-disorders-gen.; Hyperproliferation; cancer; neoplasm;
KW angiogenesis disorder; cardiovascular disease;
KW non-insulin dependent diabetes; endocrine disease;
KW gastrointestinal disease; metabolic disorder; obesity;
KW nutritional disorder; hyperlipidemia; atherosclerosis; atherogenesis;
KW hypertension; anorexia nervosa; nutritional disorder;
KW psychiatric disorder; Alzheimers disease; degeneration;
KW neurological disease; nervous system injury; neurodegenerative disease;
KW neurological disorder.
XX
OS Rattus sp.
XX
XX WO2005013901-A2.
XX
XX 17-FEB-2005.
XX
XX 30-JUL-2004; 2004WO-US025300.
XX
XX 31-JUL-2003; 2003US-0492056P.
PR 31-OCT-2003; 2003US-0516303P.
PR 19-DEC-2003; 2003US-0531596P.
PR 14-APR-2004; 2004US-0562417P.
XX
XX (ISIS-) ISIS PHARM INC.
XX
XX Esau C, Lollo B, Bennett CF, Freier SM, Griffey RH, Baker BF;
PI Vickers T, Marcussen EG, Koller E, Swayze EE, Jain R, Bhat B;
PI Peralta E;
XX
XX WPI; 2005-163123/17.
XX
XX New oligomeric compound that can hybridize with or sterically interfere
PT with nucleic acid molecules comprising or encoding small non-coding RNA
PT targets, useful for treating e.g., cancer and diabetes.
XX
XX Example 34; SEQ ID NO 1605; 854pp; English.
XX
XX The invention relates to an oligomeric compound comprising a first region
CC and a second region, where at least one region contains a modification,
CC and a portion of the oligomeric compound is targeted to a small non-
CC coding RNA target nucleic acid that is miRNA (micro-RNA), or its
CC precursor (primary-miRNA, pri-miRNA). Also included are a composition
CC comprising a first oligomeric compound and a second oligomeric compound
CC (where at least one of the oligomeric compounds contains a modification,
CC at least a portion of the first oligomeric compound is capable of
CC hybridizing with at least a portion of the second oligomeric compound,
CC and at least a portion of the first oligomeric compound is targeted to a
CC small non-coding RNA target nucleic acid), a pharmaceutical composition
CC comprising the composition cited above (and a carrier), a kit or assay
CC device comprising the composition, modulating the expression of a small
CC non-coding RNA target nucleic acid in a cell (or tissue or animal),
CC treating or preventing a disease or disorder associated with a small non-
CC coding RNA target nucleic acid, treating a condition in an animal,
CC treating or preventing a disease or disorder associated with CD3e,
CC methods of screening an oligomeric compound for an effect on miRNA
CC signaling, methods of screening a miRNA precursor for an effect in miRNA
CC signaling, methods of modulating translation (or apoptosis, conversion of

CC a precursor miRNA into miRNA, or cellular differentiation), identifying
CC an RNA transcript bound to a small non-coding RNA, arresting (or
CC delaying) entry of a cell at the G2/M phase, interfering with chromosome
CC segregation, a method of triggering apoptosis, detecting a miRNA
CC precursor, identifying a miRNA target, modulating cellular
CC differentiation, treating a condition associated with adipocyte
CC differentiation in an animal, treating/preventing a disease/disorder
CC associated with aberrant regulation of the cell cycle by miRNAs,
CC maintaining a pluripotent stem cell and identifying a small non-coding
CC RNA binding site. The oligomeric compound is targeted to a region
CC flanking a Drosha cleavage site within a pri-miRNA. It stimulates an
CC increase in expression of a pri-miRNA. The compounds and compositions are
CC useful for treating a disease or disorder resulting from chromosomal non-
CC disjunction, altered methylation, acetylation, or pseudouridylation state
CC of chromosomes, such as a hyperproliferative condition (e.g. cancer,
CC neoplasia or angiogenesis), diabetes (Type 2 diabetes), obesity,
CC hyperlipidemia, atherosclerosis, atherogenesis, hypertension, anorexia,
CC Alzheimer's disease, a central nervous system injury or neurodegenerative
CC disorder. The present sequence is a primary miRNA of the invention.
XX
SQ Sequence 95 BP; 18 A; 22 C; 32 G; 0 T; 23 U; 0 Other;
Query Match 71.2%; Score 17.8; DB 14; Length 95;
Best Local Similarity 71.4%; Pred. No. 1.3e+02;
Matches 15; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
QY 5 AAGAGCGCAGACTGCTCTTCC 25
Db 61 AAGAGCACAGCUGCUCUCC 81
RESULT 37
ABS69109
ID ABS69109 standard; DNA; 331 BP.
XX
AC ABS69109;
XX
DT 21-NOV-2002 (first entry)
XX
DE Novel murine polynucleotide isolated using gene trap technology #172.
XX
XX Mouse; gene trapped sequence; GTS; functional genomic analysis;
KW phage display system; gene chip; temporal gene expression;
KW tissue specific gene expression; antisense inhibition; gene targeting;
KW development disorder; cell differentiation disorder; aging; cancer;
KW autoimmune disease; lupus; inflammatory disorder; skin disorder;
KW degenerative disorder; ds.
XX
XX Mus musculus.
XX
XX US2002102543-A1.
XX
XX 01-AUG-2002.
XX
XX 30-NOV-2000; 2000US-00728445.
PF
XX 01-DEC-1999; 99US-0168358P.
PR
XX (FRIE/) FRIEDRICH G.
PA (ZAMB/) ZAMBROWICZ B.
PA (SAND/) SANDS A T.
XX
XX Friedrich G, Zambrowicz B, Sands AT;
PI WPI; 2002-690598/74.
DR
XX Novel murine polynucleotides that individually identify novel genes into
PT which a retroviral gene trap vector has integrated, useful in genomic
PT analysis and in discovery, development of therapeutic and diagnostic
PT agents.
XX
XX Claim 1; Page 75; 296pp; English.
XX

CC The invention describes an isolated murine polynucleotide (I) comprising
CC a contiguous stretch of at least 60 nucleotides of one of 265-677
CC nucleotide 891 OMNIBANK gene trapped sequences (GTSs) (S), given in the
CC specification. The novel genes and cells are useful in functional genomic
CC analysis and in the discovery and development of new therapeutic and
CC diagnostic agents and methods. (I) is useful for identifying the coding
CC regions of the murine genome, to isolate cDNAs, genomic clones, or full-
CC length genes/polynucleotides or homologues, heterologues, paralogues, or
CC orthologues that are capable of hybridising to one or more of the GTSs
CC under stringent conditions. (I) can be incorporated into a phage display
CC system that can be used to screen for proteins, or other ligands, that
CC are capable of binding an amino acid sequence encoded by an
CC oligonucleotide or polynucleotide sequence in at least one of the TS
CC sequences. (I) is useful in addressable arrays, such as gene chips, to
CC identify and characterise temporal and tissue specific gene expression,
CC to identify the gene of interest from many sources and for genetic
CC manipulations such as antisense inhibition and gene targeting. Decreasing
CC the level of expression of (I) and/or down regulating the activity of
CC peptides or proteins encoded by (I) is useful for treating development
CC and cell differentiation disorders, aging, cancer, autoimmune disease,
CC lupus, inflammatory disorders, skin disorders and degenerative disorders.
CC This sequence represents a murine cDNA isolated using gene trap
CC technology
XX
SQ Sequence 331 BP; 70 A; 79 C; 82 G; 92 T; 0 U; 8 Other;

Query Match 71.2%; Score 17.8; DB 6; Length 331;
Best Local Similarity 90.5%; Pred. No. 1.7e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 AAGAGCGCAGACTGGTCTTCC 25
||||| ||||| ||||| |||||
Db 184 AAGAGCACAGACTGCTTCC 204

RESULT 38
ADO36057
ID ADO36057 standard; DNA; 3486 BP.
XX
AC ADO36057;
XX
XX 26-AUG-2004 (first entry)
XX
DE Novel mouse gene sequence #730.
XX
XX mouse; murine; cancer; psoriasis; ulcerative colitis; inflammation;
KW ischaemic heart disease; thrombosis; immune disorder; bacterial disorder;
KW viral disorder; ds; gene.
XX
OS Mus sp.
XX
XX WO2004046310-A2.
XX
XX 03-JUN-2004.
XX
XX 24-OCT-2003; 2003WO-US033948.
XX
XX 15-NOV-2002; 2002US-0426916P.
PR 04-DEC-2002; 2002US-0431158P.
PR 05-DEC-2002; 2002US-0431445P.
PR 05-DEC-2002; 2002US-0431606P.
PR 09-JUN-2003; 2003US-0476621P.
PR 09-JUN-2003; 2003US-0476632P.
PR 08-JUL-2003; 2003US-0485217P.
PR 08-AUG-2003; 2003US-0485359P.
PR 08-AUG-2003; 2003US-0493332P.
PR 08-AUG-2003; 2003US-0493356P.

(FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
XX Williams LT, Chu K, Lee E, Hestir K, Hayashizaki Y, Kamiya M;
XX WPI; 2004-431966/40.

XX New mouse nucleic acid molecules and polypeptides, useful for treating
PT cancer, psoriasis, ulcerative colitis, inflammation, ischemic heart
PT disease or thrombosis.
XX
XX Claim 1; SEQ ID NO 730; 263pp; English.
XX
XX The invention comprises 744 novel mouse DNA sequences (genes). The DNA
XX sequences of the invention are useful for treating cancer, psoriasis,
CC ulcerative colitis, inflammation, ischaemic heart disease, thrombosis,
CC immune disorders, bacterial disorders and viral disorders. The present
CC nucleic acid represents a mouse DNA sequence of the invention. NOTE: The
CC present DNA sequence is not shown in the specification, but has been
CC retrieved from the WIPO website.

XX
SQ Sequence 3486 BP; 836 A; 924 C; 796 G; 929 T; 0 U; 1 Other;
Query Match 71.2%; Score 17.8; DB 12; Length 3486;
Best Local Similarity 90.5%; Pred. No. 2.6e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 AAGAGCGCAGACTGGTCTTCC 25
||||| ||||| ||||| |||||
Db 2819 AAGAGCACAGACTGCTTCC 2839

RESULT 39
ADO36057/C
ID ADO36057 standard; DNA; 3486 BP.
XX

AC ADO36057;
XX
XX 26-AUG-2004 (first entry)
XX
DE Novel mouse gene sequence #730.

XX mouse; murine; cancer; psoriasis; ulcerative colitis; inflammation;
KW ischaemic heart disease; thrombosis; immune disorder; bacterial disorder;
KW viral disorder; ds; gene.
XX
XX Mus sp.

XX WO2004046310-A2.
XX
XX 03-JUN-2004.
XX
XX 24-OCT-2003; 2003WO-US033948.

XX
XX 15-NOV-2002; 2002US-0426916P.
PR 04-DEC-2002; 2002US-0431158P.
PR 05-DEC-2002; 2002US-0431445P.
PR 05-DEC-2002; 2002US-0431606P.
PR 09-JUN-2003; 2003US-0476621P.
PR 09-JUN-2003; 2003US-0476632P.
PR 08-JUL-2003; 2003US-0485217P.
PR 08-JUL-2003; 2003US-0485359P.
PR 08-AUG-2003; 2003US-0493332P.
PR 08-AUG-2003; 2003US-0493356P.

(FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
XX Williams LT, Chu K, Lee E, Hestir K, Hayashizaki Y, Kamiya M;
XX WPI; 2004-431966/40.

XX New mouse nucleic acid molecules and polypeptides, useful for treating
PT cancer, psoriasis, ulcerative colitis, inflammation, ischemic heart
PT disease or thrombosis.
XX
XX Claim 1; SEQ ID NO 730; 263pp; English.

XX The invention comprises 744 novel mouse DNA sequences (genes). The DNA
XX sequences of the invention are useful for treating cancer, psoriasis,
CC disease or thrombosis.

CC ulcerative colitis, inflammation, ischaemic heart disease, thrombosis,
CC immune disorders, bacterial disorders and viral disorders. The present
CC nucleic acid represents a mouse DNA sequence of the invention. NOTE: The
CC present DNA sequence is not shown in the specification, but has been
CC retrieved from the WIPO website.

XX
SQ Sequence 3486 BP; 836 A; 924 C; 796 G; 929 T; 0 U; 1 Other;
Query Match 71.2%; Score 17.8; DB 12; Length 3486;
Best Local Similarity 90.5%; Pred. No. 2.6e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 AAGAGCGCAGACTGCTCTTCC 25
ADB11347
|||||
Db 164 AAGAGCGCAGACTGCTCTTCC 144

RESULT 40
ADB11347
ID ADB11347 standard; DNA; 4333 BP.
XX
AC ADB11347;
XX
DT 04-DEC-2003 (first entry)
XX
DE Murine downstream homology arm inserted into the ROSA26 targeting vector.

XX ds; phic31 integrase; ROSA26 promoter; codon optimised;
KW site specific recombinase; SSR; gene function; disease model;
KW gene therapy; transgenic; C31-Int; mouse; murine; homology arm.
XX
OS Mus sp.

XX WO2003066867-A2.
PN
XX 14-AUG-2003.
XX
XX 05-FEB-2003; 2003WO-EP001122.
PF
XX 06-FEB-2002; 2002US-0354741P.

XX (ARTE-) ARTEMIS PHARM GMBH.
PA
XX Andreas S, Faust N;
PI
XX WPI; 2003-663599/c2.

XX New genetically engineered nucleic acid molecule, useful for preparing an
PT agent for recombining a DNA molecule containing phic31 integrase
PT recognition sequences in a eukaryotic cell, a vertebrate or transgenic
PT organism.

XX Example 3; Page 71-73; 87pp; English.

XX This invention relates to novel genetically engineered nucleic acid
CC molecules encoding phic31 integrase (C31-Int), which has been codon
CC optimised for expression in eukaryotic host cells. The phic31 integrase
CC is a site specific recombinase (SSR) that catalyzes recombination between
CC two phic31 recognition sequences. The introduction of silent mutations
CC into the coding sequence changes the given codon to one that is most
CC frequently used in the respective host, which in turn alters expression
CC levels. Accordingly, using this ability to generate controlled and
CC permanent modifications in eukaryotic genomes has various research
CC applications including the study of gene function and the creation of
CC disease models, as well as gene therapy for medical applications, and the
CC design of economically important animals and crops. Furthermore, the
CC phic31 integrase of the invention is useful for preparing an agent for
CC recombining a DNA molecule containing phic31 integrase recognition
CC sequences in a eukaryotic cell, a vertebrate or transgenic organism. This
CC polynucleotide sequence is the murine downstream homology arm located in
CC the ROSA26 promoter targeting vectors and required for homologous
CC recombination in ES cells, used in an exemplification of the invention.

SQ Sequence 4333 BP; 1177 A; 814 C; 860 G; 1482 T; 0 U; 0 Other;
Query Match 71.2%; Score 17.8; DB 10; Length 4333;
Best Local Similarity 90.5%; Pred. No. 2.8e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 AAGAGCGCAGACTGCTCTTCC 25
ADB11347
|||||
Db 1308 AAGAGCACAGACTGCTCTTCC 1328

RESULT 41
ADQ76467
ID ADQ76467 standard; DNA; 4333 BP.
XX
AC ADQ76467;

XX 07-OCT-2004 (first entry)
XX
XX Nucleotide sequence of a Rosa26 targeting vector 3' arm.
DE targeted transgenesis; Rosa26 locus; homologous recombination;
KW transgenic eukaryotic cell; gene function study; drug development;
KW disease model; ss.
XX
OS Mus musculus.

XX EP1439234-A1.
PN
XX 21-JUL-2004.
PD
XX 08-JAN-2003; 2003EP-00000249.

XX 08-JAN-2003; 2003EP-00000249.
XX
XX (ARTE-) ARTEMIS PHARM GMBH.
PA
XX Schwenk F, Seibler J, Faust N, Kuehn R;
PI
XX WPI; 2004-509160/49.

XX Generating transgenic eukaryotic cells having a modified Rosa26 locus
PT comprises introducing a functional DNA sequence into the Rosa26 locus of
PT starting eukaryotic cells by homologous recombination with a targeting
PT vector.
XX
PS Disclosure; SEQ ID NO 5; 26pp; English.

XX The specification describes a method for targeted transgenesis, which
CC generates transgenic eukaryotic cells having a modified Rosa26 locus. The
CC method comprises introducing a functional DNA sequence into the Rosa26
CC locus of starting eukaryotic cells by homologous recombination with a
CC targeting vector comprising the functional DNA sequence flanked by DNA
CC sequences homologous to the Rosa26 locus. Rosa26 supports the
CC preservation of inherent activity of heterologous promoters inserted
CC through homologous recombination into the locus. The method of the
CC invention is useful for generating transgenic eukaryotic cells. The
CC eukaryotic cells are useful for gene function studies, drug development
CC and as disease models. ADQ76466-ADQ76467 represent 5' and 3' arms of a
CC Rosa26 targeting vector. These two sequences enable homologous
CC recombination into the Rosa locus.

XX Sequence 4333 BP; 1177 A; 814 C; 860 G; 1482 T; 0 U; 0 Other;
Query Match 71.2%; Score 17.8; DB 12; Length 4333;
Best Local Similarity 90.5%; Pred. No. 2.8e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 AAGAGCGCAGACTGCTCTTCC 25
ADB11347
|||||
Db 1308 AAGAGCACAGACTGCTCTTCC 1328


```
RESULT 42
ADQ82251
ID ADQ82251 standard; DNA; 4333 BP.
XX
XX ADQ82251;
AC
XX 21-OCT-2004 (first entry)
DT
XX
XX Rosa26 locus 3' arm.
DE
XX transgenic eukaryotic cell; Rosa26 locus; gene function study;
KW drug development; ds.
XX
XX Mus sp.
OS
XX
XX Key Location/Qualifiers
FH 1. .1616
FT promoter
FT /*tag= a
FT /*note= "CAGGS-promoter"
FT polyA_site 3921. .4099
FT /*tag= b
FT /*note= "CreTR site"
XX
XX WO2004063381-A1.
XX
XX 29-JUL-2004.
XX
XX 08-JAN-2004; 2004WO-EP000065.
XX
XX 08-JAN-2003; 2003EP-00000249.
XX
XX 10-JAN-2003; 2003US-0439367P.
XX
XX (ARTE-) ARTEMIS PHARM GMBH.
XX
XX Schwenk F, Seibler J, Faust N, Kuehn R;
PI
XX WPI; 2004-544090/52.
XX
XX Generating transgenic eukaryotic cells and animals having a modified
PT Rosa26 locus by introducing a functional DNA sequence into the Rosa26
PT locus, useful in gene function studies and drug development.
XX
XX Claim 6; SEQ ID NO 5; 53pp; English.
XX
XX The present invention relates to generating transgenic eukaryotic cells
CC having a modified Rosa26 locus comprising introducing a functional DNA
CC sequence into the Rosa26 locus of starting eukaryotic cells, where the
CC functional DNA sequence is a gene expression cassette comprising a gene
CC of interest operatively linked to a promoter or is a DNA sequence which
CC can be converted into such a gene expression cassette. The methods and
CC compositions of the present invention are useful in targeted transgenesis
CC using the Rosa26 locus. The transgenic cells and organisms are useful for
CC gene function studies, drug development and as disease model animals. The
CC present sequence represents the Rosa26 locus 3' arm.
XX
XX Sequence 4333 BP; 1177 A; 814 C; 860 G; 1482 T; 0 U; 0 Other;
SQ
Query Match 71.2%; Score 17.8; DB 13; Length 4333;
Best Local Similarity 90.5%; Pred. No. 2.8e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 5 AAGAGCGCAGACTGGTCTTCC 25
||||| ||||| |||||
Db 1308 AAGAGCACAGACTGCTCTTCC 1328
RESULT 43
ADQ84031
ID ADR04031 standard; DNA; 4333 BP.
XX
XX ADR04031;
AC
XX 21-OCT-2004 (first entry)
DT
XX
XX Murine Rosa26 gene 3' arm.
DE
XX ds; mouse; Rosa26; targeted homologous recombination; gene; transgenic.
XX
XX Mus musculus.
XX
XX EP1445320-A1.
XX
XX 11-AUG-2004.
XX
XX 05-FEB-2003; 2003EP-00002424.
XX
XX 05-FEB-2003; 2003EP-00002424.
XX
XX (ARTE-) ARTEMIS PHARM GMBH.
XX
XX Kauselmann G, Zevnik B, Seibler J, Kern H;
XX
XX WPI; 2004-563471/55.
XX
XX Novel vector comprising homologous targeting cassette harboring
PT functional DNA sequence, and expression cassette harboring DNA sequence
PT coding for first detectable marker, useful for targeted homologous
PT recombination of eukaryotic cells.
XX
XX Example 1; SEQ ID NO 5; 46pp; English.
XX
XX The present invention relates to a vector for targeted homologous
CC recombination of eukaryotic cells, comprising a homologous targeting
CC cassette harbouring a functional DNA sequence, and an expression cassette
CC harbouring a functional DNA sequence coding for a first detectable
CC marker, with the expression cassette being connected with the homologous
CC targeting cassette so as to allow distinction between targeted and non-
CC targeted cells. The vector is useful for homologous recombination of
CC eukaryotic cells, which is useful for preparing transgenic tissues,
CC organs and/or multi-cell organisms. The transgenic multi-cell organism is
CC a non-human mammal, most preferably mouse. The present sequence is a
CC fragment of the murine Rosa26 gene used in the exemplification of the
CC invention.
XX
XX Sequence 4333 BP; 1177 A; 814 C; 860 G; 1482 T; 0 U; 0 Other;
SQ
Query Match 71.2%; Score 17.8; DB 13; Length 4333;
Best Local Similarity 90.5%; Pred. No. 2.8e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 5 AAGAGCGCAGACTGGTCTTCC 25
||||| ||||| |||||
Db 1308 AAGAGCACAGACTGCTCTTCC 1328
RESULT 44
ADT71363
ID ADT71363 standard; DNA; 4333 BP.
XX
XX ADT71363;
AC
XX 16-DEC-2004 (first entry)
DT
XX 3' arm Rosa26.
DE
XX Rosa26; ds; mouse; homologous recombination; detectable marker;
KW fluorescence marker; positive selection marker; first detectable marker;
KW RRS; recombinase recognition site; optical selection;
KW second detectable marker.
XX
XX Mus musculus.
XX
XX WO2004070040-A1.
XX
XX 19-AUG-2004.
XX
XX
```


OS unidentified adenovirus.
 XX Mus sp.
 XX
 PN WO2003066867-A2.
 XX
 PD 14-AUG-2003.
 XX
 PF 05-FEB-2003; 2003WO-EP001122.
 XX
 PR 06-FEB-2002; 2002US-0354741P.
 XX
 XX (ARTE-) ARTEMIS PHARM GMBH.
 XX
 XX Andreas S, Faust N;
 XX
 PI WPI; 2003-663599/62.
 DR
 XX
 XX New genetically engineered nucleic acid molecule, useful for preparing an
 PT agent for recombining a DNA molecule containing phic31 integrase
 PT recognition sequences in a eukaryotic cell, a vertebrate or transgenic
 PT organism.
 XX
 XX Example 3; Page 73-79; 87pp; English.
 PS
 XX This invention relates to novel genetically engineered nucleic acid
 CC molecules encoding phic31 integrase (C31-Int), which has been codon
 CC optimised for expression in eukaryotic host cells. The phic31 integrase
 CC is a site specific recombinase (SSR) that catalyzes recombination between
 CC two phic31 recognition sequences. The introduction of silent mutations
 CC into the coding sequence changes the given codon to one that is most
 CC frequently used in the respective host, which in turn alters expression
 CC levels. Accordingly, using this ability to generate controlled and
 CC permanent modifications in eukaryotic genomes has various research
 CC applications including the study of gene function and the creation of
 CC disease models, as well as gene therapy for medical applications, and the
 CC design of economically important animals and crops. Furthermore, the
 CC phic31 integrase of the invention is useful for preparing an agent for
 CC recombining a DNA molecule containing phic31 integrase recognition
 CC sequences in a eukaryotic cell, a vertebrate or transgenic organism. This
 CC polynucleotide sequence is the pRosa12 vector comprising an adenoviral
 CC splice acceptor site inserted between the two ROSA26 promoter homology
 CC arms, used in an exemplification of the invention.
 XX
 SQ Sequence 10491 BP; 2499 A; 2433 C; 2608 G; 2951 T; 0 U; 0 Other;
 Query Match 71.2%; Score 17.8; DB 10; Length 10491;
 Best Local Similarity 90.5%; Pred. No. 3.3e+02;
 Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 5 AAGAGCGCAGACTGGTCTTCC 25
 DB 4954 AAGAGCACAGACTGCTCTTCC 4974
 RESULT 47
 AAS01080
 ID AAS01080 standard; DNA; 11176 BP.
 XX
 AC AAS01080;
 XX
 DT 29-MAY-2001 (first entry)
 XX
 DE Mouse FosB DNA.
 XX
 KW Mouse; phosphoglycerate kinase 1 promoter; targeting construct;
 KW transgenic animal; reporter expression cassette; luciferase; vitronectin;
 KW FosB; galactin 3; VSGF; VEGFR; Tie2; selectable marker;
 KW neomycin phosphotransferase II; chloramphenicol acetyltransferase;
 KW thymidine kinase; PCR primer; ds.
 XX
 OS Mus sp.
 XX
 PN WO200118225-A1.

XX 15-MAR-2001.
 XX
 XX 16-DEC-1999; 99WO-US030078.
 XX
 PR 03-SEP-1999; 99US-0152522P.
 XX
 XX (XENO-) XENOGEN CORP.
 XX
 XX Zhang N;
 XX
 XX WPI; 2001-203085/20.
 XX
 XX Targeting constructs and methods of using them for creating transgenic
 PT animals.
 PT
 XX Disclosure; Fig 4B; 96pp; English.
 PS
 XX The present sequence for mouse FosB DNA is used to generate a targeting
 CC cassette and vector in the invention of novel targeting constructs and
 CC methods of using them for creating transgenic animals in which at least 1
 CC single-copy, non-essential gene is replaced with a reporter expression
 CC cassette (e.g. a luciferase gene linked to a promoter heterologous to the
 CC single-copy, non-essential gene). The single-copy, non-essential gene may
 CC be selected from vitronectin (Vn), fosB and galactin 3. The promoter
 CC element may be selected from VEGF, VEGFR and Tie2. The constructs may
 CC contain a selectable marker e.g. neomycin phosphotransferase II,
 CC chloramphenicol acetyltransferase and thymidine kinase. The constructs
 CC are useful for creating transgenic animals in which at least 1 single-
 CC copy, non-essential gene is replaced with a reporter expression cassette.
 CC Methods of using these animals are also described
 XX
 SQ Sequence 11176 BP; 2570 A; 2863 C; 2801 G; 2941 T; 0 U; 1 Other;
 Query Match 71.2%; Score 17.8; DB 4; Length 11176;
 Best Local Similarity 90.5%; Pred. No. 3.3e+02;
 Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 5 AAGAGCGCAGACTGGTCTTCC 25
 DB 10526 AAGAGCACAGACTGCTCTTCC 10546
 RESULT 48
 AAD22528
 ID AAD22528 standard; DNA; 11176 BP.
 XX
 AC AAD22528;
 XX
 DT 12-FEB-2002 (first entry)
 XX
 DE Mouse FosB DNA.
 XX
 KW Transcription control element; angiogenesis; tumorigenesis; VEGFR-2;
 KW transgenic animal; vascular endothelial growth factor; VEGF; Tie2 gene;
 KW VEGF receptor 2; PCR primer; FosB gene; mouse; ss.
 XX
 OS Mus sp.
 XX
 XX US2001037016-A1.
 PN
 XX
 PD 01-NOV-2001.
 XX
 PF 15-DEC-2000; 2000US-00738968.
 XX
 PR 16-DEC-1999; 99US-00465978.
 XX
 XX (NING/) NING Z.
 XX (CONT/) CONTAG P R.
 XX (PURC/) PURCHIO A F.
 XX
 XX Ning Z, Contag PR, Purchio AF;
 PI
 XX

DR WPI; 2002-017156/02.
XX Polynucleotide transcription control elements derived from vascular
PT endothelial growth factor, and receptors useful for identifying
PT modulators useful for modulating tumorigenesis.
XX Example 1; Page 45-50; 85pp; English.
XX The invention relates to transcription control elements including
CC promoters derived from mouse vascular endothelial growth factor (VEGF),
CC VEGF receptor 2 (VEGFR-2) and Tie2 gene loci. The invention also relates
CC to expression cassettes and vector constructs incorporating promoters
CC useful for generating transformed cells and transgenic animals and
CC methods for screening therapeutic compounds that modulate angiogenesis
CC and tumorigenesis. The present sequence is mouse FosB DNA used in the
CC construction of vectors containing VEGF, VEGFR-2 and Tie2 expression
CC cassettes
XX
SQ Sequence 11176 BP; 2570 A; 2863 C; 2801 G; 2941 T; 0 U; 1 Other;
Query Match 71.2%; Score 17.8; DB 6; Length 11176;
Best Local Similarity 90.5%; Pred. No. 3.3e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 5 AAGAGCGCAGACTGCTCTTCC 25
DB 10526 AAGAGCAGACAGCTGCTCTTCC 10546
RESULT 49
ABT08201
ID ABT08201 standard; DNA; 11784 BP.
AC ABT08201;
XX 28-NOV-2002 (first entry)
XX Modified ROSA26 locus.
XX Fusion protein; recombinase domain; signal peptide domain; gene function;
KW nuclear import; recombinase recognition sequence; transgenic organism;
KW C31-Int recombination system; site-specific integration; gene therapy;
KW ds.
XX Unidentified.
OS Synthetic.
XX WO200238613-A2.
PN 16-MAY-2002.
PD 09-NOV-2001; 2001WO-EP012975.
PF 10-NOV-2000; 2000EP-00124629.
PR 17-APR-2001; 2001EP-00109543.
PR 13-AUG-2001; 2001US-0311876P.
XX (ARTE-) ARTEMIS PHARM GMBH.
PA Kuehn R, Felder S, Schwenk F, Kueter-Luks B, Faust N;
XX WPI; 2002-519298/55.
XX Novel fusion protein useful for recombining DNA molecules in eukaryotic
PT cells has recombinase protein which is linked to signal peptide domain
PT which directs nuclear import of fusion protein in eukaryotic cells.
XX Example 4; Page 135-138; 150pp; English.
XX The invention comprises the amino acid and coding sequences of fusion
CC proteins that contain a recombinase domain and a signal peptide domain
CC which directs nuclear import of the fusion protein in eukaryotic cells.
CC The fusion proteins of the invention are useful for recombining the DNA

CC molecules of cells or organisms containing recombinase recognition
CC sequences for the recombinase domain of the fusion proteins. The fusion
CC proteins of the invention are useful for studying gene function at
CC various developmental stages and for the creation of transgenic
CC organisms. The C31-Int recombination system of the invention can be used
CC for the site-specific integration of foreign DNA into the genome of
CC mammalian cells (e.g. for gene therapy). The present DNA sequence
CC represents an oligonucleotide that was used in the invention
XX
SQ Sequence 11784 BP; 2796 A; 2787 C; 3027 G; 3174 T; 0 U; 0 Other;
Query Match 71.2%; Score 17.8; DB 6; Length 11784;
Best Local Similarity 90.5%; Pred. No. 3.3e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 5 AAGAGCGCAGACTGCTCTTCC 25
DB 8760 AAGAGCAGACAGCTGCTCTTCC 8780
RESULT 50
ADB81350
ID ADB81350 standard; DNA; 11784 BP.
XX AC ADB81350;
XX 04-DEC-2003 (first entry)
XX Bacteriophage C31 substrate reporter DNA.
XX ROSA26 promoter; ds; phiC31 integrase; PROSA12;
KW site specific recombinase; SSR; gene function; disease model;
KW gene therapy; transgenic; C31-Int.
XX Bacteriophage phi-C31.
PN WO2003066867-A2.
XX 14-AUG-2003.
XX 05-FEB-2003; 2003WO-EP001122.
XX 06-FEB-2002; 2002US-0354741P.
XX (ARTE-) ARTEMIS PHARM GMBH.
XX Andreas S, Faust N;
XX WPI; 2003-663599/62.
XX New genetically engineered nucleic acid molecule, useful for preparing an
PT agent for recombining a DNA molecule containing phiC31 integrase
PT recognition sequences in a eukaryotic cell, a vertebrate or transgenic
PT organism.
XX Example 3; Page 79-86; 87pp; English.
XX This invention relates to novel genetically engineered nucleic acid
CC molecules encoding phiC31 integrase (C31-int), which has been codon
CC optimised for expression in eukaryotic host cells. The phiC31 integrase
CC is a site specific recombinase (SSR) that catalyzes recombination between
CC two phiC31 recognition sequences. The introduction of silent mutations
CC into the coding sequence changes the given codon to one that is most
CC frequently used in the respective host, which in turn alters expression
CC levels. Accordingly, using this ability to generate controlled and
CC permanent modifications in eukaryotic genomes has various research
CC applications including the study of gene function and the creation of
CC disease models, as well as gene therapy for medical applications, and the
CC design of economically important animals and crops. Furthermore, the
CC phiC31 integrase of the invention is useful for preparing an agent for
CC recombining a DNA molecule containing phiC31 integrase recognition
CC sequences in a eukaryotic cell, a vertebrate or transgenic organism. This
CC polynucleotide sequence is the bacteriophage phi-31 substrate reporter

CC DNA located in the ROSA26 locus of heterozygous mice, used in an
CC exemplification of the invention.
XX
SQ Sequence 11784 BP; 2796 A; 2787 C; 3027 G; 3174 T; 0 U; 0 Other;
Query Match 71.2%; Score 17.8; DB 10; Length 11784;
Best Local Similarity 90.5%; Pred. NO. 3.3e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 5 AGAGCGCAGACTGGTCTTCC 25
Db 8760 AGAGCACAGACTGCTCTTCC 8780

Search completed: February 3, 2006, 21:57:20
Job time : 210.111 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 3, 2006, 14:38:25 ; Search time 2951.67 Seconds
(without alignments)
396.277 Million cell updates/sec

Title: US-10-719-900-35
Perfect score: 25
Sequence: 1 aaaaagagcgagactggtcttcc 25

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 150 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	23.4	93.6	655	1	BB066943
2	23.4	93.6	2333	4	AK033191 Mus muscu
3	21.8	87.2	346	5	BY306264 BY306264
4	21.8	87.2	760	10	AG405809 Mus muscu
5	21.8	87.2	938	5	BQ925815 AGENCOURT
6	20.2	80.8	212	2	BB371564 BB371564
7	20.2	80.8	249	6	CA504477 UI-R-P81-
8	20.2	80.8	330	9	BZ192915 CH230-248
9	20.2	80.8	347	1	AU018134 AU018134
10	20.2	80.8	352	9	AZ947222 2M0209B21
11	20.2	80.8	358	2	BG071090 H3094F01-
12	20.2	80.8	363	9	BH315230 CH230-159
13	20.2	80.8	371	2	BG083952 H3094F01-
14	20.2	80.8	393	9	BH344887 CH230-183
15	20.2	80.8	400	7	CM669196 A0876H02-
16	20.2	80.8	417	1	AW535711 UI-R-B50-
17	20.2	80.8	420	9	BZ207783 CH230-475
18	20.2	80.8	430	5	BY567189 BY567189
19	20.2	80.8	443	3	BP417531 BP417531
20	20.2	80.8	464	9	AZ960237 2M0228P09
21	20.2	80.8	480	9	BH260028 CH230-113
22	20.2	80.8	481	5	BY566957 BY566957

C 23	20.2	80.8	482	3	BP775662
C 24	20.2	80.8	506	3	BI301386
C 25	20.2	80.8	509	9	AZ875623 2M0190D13
C 26	20.2	80.8	528	2	BB756913 BB756913
C 27	20.2	80.8	536	9	BH349233 CH230-81B
C 28	20.2	80.8	574	2	BG088328 H3151G04-
C 29	20.2	80.8	577	9	AZ288000 RPCI-23-1
C 30	20.2	80.8	586	9	BH261566 CH230-90J
C 31	20.2	80.8	607	9	AZ357207 1M0098D08
C 32	20.2	80.8	625	7	CN542163 UI-R-EAO-
C 33	20.2	80.8	632	9	BZ116863 CH230-510
C 34	20.2	80.8	702	10	AG282644 Mus muscu
C 35	20.2	80.8	708	9	BZ196833 CH230-269
C 36	20.2	80.8	714	10	AG562735 Mus muscu
C 37	20.2	80.8	726	2	BG298680 602396890
C 38	20.2	80.8	736	9	AZ430769 1M0215G22
C 39	20.2	80.8	779	9	BZ173903 CH230-493
C 40	20.2	80.8	789	9	BH292979 CH230-92B
C 41	20.2	80.8	809	9	BZ277088 CH230-308
C 42	20.2	80.8	816	10	AG295893 Mus muscu
C 43	20.2	80.8	831	5	BQ959424 AGENCOURT
C 44	20.2	80.8	1052	9	BZ275313 CH230-450
C 45	19.8	79.2	401	1	AA138597 mg92C11.r
C 46	19.8	79.2	868	7	CO034709 ESP813093
C 47	19.4	77.6	159	1	AA590658 vm24G12.r
C 48	19.4	77.6	277	1	AA823597 vw38a05.r
C 49	19.4	77.6	292	11	CR109101 Reverse s
C 50	19.4	77.6	343	2	BE197804 us77e02.x
C 51	19.4	77.6	347	2	BG094937 mac25h02.x
C 52	19.4	77.6	356	5	BY210447 BY210447
C 53	19.4	77.6	365	5	BY209501 BY209501
C 54	19.4	77.6	375	1	AI413095 md92C07.r
C 55	19.4	77.6	382	8	W62939 md92C07.r1
C 56	19.4	77.6	389	8	W48161 mc8606.r1
C 57	19.4	77.6	427	5	BY545965 BY545965
C 58	19.4	77.6	436	1	AA822961 vw39a05.r
C 59	19.4	77.6	443	1	AI430486 md32C07.y
C 60	19.4	77.6	467	5	BY256574 BY256574
C 61	19.4	77.6	475	1	AA690208 vt78G11.s
C 62	19.4	77.6	477	1	AI427277 mc86e06.y
C 63	19.4	77.6	487	9	AZ017695 RPCI-23-2
C 64	19.4	77.6	494	9	AZ403580 1M0171007
C 65	19.4	77.6	516	1	AA164091 mr23D07.r
C 66	19.4	77.6	543	4	AK020460 Mus muscu
C 67	19.4	77.6	558	5	BY721720 BY721720
C 68	19.4	77.6	564	2	BG609520 323403 MA
C 69	19.4	77.6	597	11	CR043665 Reverse s
C 70	19.4	77.6	604	7	CN664053 A0808B01-
C 71	19.4	77.6	605	9	AZ403386 1M0171007
C 72	19.4	77.6	609	5	BQ604998 MI-P-CPI-
C 73	19.4	77.6	626	2	BB664023 BB664023
C 74	19.4	77.6	635	11	CR258177 Forward s
C 75	19.4	77.6	642	5	BY718759 BY718759
C 76	19.4	77.6	648	6	CF171469 B0842H03-
C 77	19.4	77.6	661	9	AZ839688 2M0135K23
C 78	19.4	77.6	686	5	BX669334 BX669334
C 79	19.4	77.6	686	7	CJ017268 CJ017268
C 80	19.4	77.6	718	7	CJ038423 CJ038423
C 81	19.4	77.6	815	3	BP172852 BP172852
C 82	19.4	77.6	848	11	CR192365 Reverse s
C 83	19.4	77.6	862	11	CR213048 Reverse s
C 84	19.4	77.6	1056	8	W83215 mf23C02.r1
C 85	19.4	77.6	1102	4	AK078159 Mus muscu
C 86	19.4	77.6	1761	4	AK087053 Mus muscu
C 87	19.2	76.8	193	3	BQ166077 WHE0808.D
C 88	19.2	76.8	390	9	AZ108024 RPCI-23-2
C 89	19.2	76.8	406	9	BZ197313 CH230-416
C 90	19.2	76.8	477	9	AZ613766 1M0442113
C 91	19.2	76.8	548	7	CF898666 A0242F05-
C 92	19.2	76.8	579	9	AQ487165 RPCI-11-2
C 93	19.2	76.8	599	9	AQ486429 RPCI-11-2
C 94	19.2	76.8	601	11	CR180552 Reverse s
C 95	19.2	76.8	607	9	BH294574 CH230-104

Db	45	AAAAAAGAGCACTGACTGCTCTTCC	21
RESULT 4	AG405809	760 bp	DNA linear GSS 21-DEC-2004
LOCUS	Mus musculus molossinus DNA, clone:MSMg01-264D02.TJ, genomic survey		
DEFINITION	sequence.		
ACCESSION	AG405809		
VERSION	AG405809.1	GI:48048495	
KEYWORDS	GSS.		
SOURCE	Mus musculus molossinus (Japanese wild mouse)		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
REFERENCE	1		
AUTHORS	Abe, K., Noguchi, H., Tagawa, K., Yuzuriha, M., Toyoda, A., Kojima, T., Ezawa, K., Saitou, N., Hattori, M., Sakaki, Y., Moriawaki, K. and Shiroishi, T.		
TITLE	Contribution of Asian mouse subspecies Mus musculus molossinus to genomic constitution of strain C57BL/6J, as defined by BAC-end sequence-SNP analysis		
JOURNAL	Genome Res. 14 (12), 2439-2447 (2004)		
PUBMED	15574823		
REFERENCE	2 (bases 1 to 760)		
AUTHORS	Hattori, M., Toyoda, A., Noguchi, H., Kojima, T. and Sakaki, Y.		
TITLE	Direct Submission		
JOURNAL	Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa, 230-0045, Japan (E-mail: hattori@gsc.riken.jp URL: http://hgp.gsc.riken.go.jp/, Tel: 81-45-503-9111, Fax: 81-45-503-9170)		
COMMENT	Clones are derived from the mouse BAC library MSMg01. For BAC library availability, please contact Kuniya Abe (abe@tc.riken.jp). Tsukuba Institute, Bio Resource Center, The Institute of Physical and Chemical Research (RIKEN) 3-1-1 Koyadai, Tsukuba 305-0074 Japan Phone: 81-298-36-9189, fax: 81-298-36-9199 e-mail: abe@tc.riken.jp		
PRIMERS	Sequencing : TJ		
LIBRARY	: DBACE3.6		
Vector	: BclRI		
R.Site 1	: BclRI		
R.Site 2	: BclRI		
FEATURES	Location/Qualifiers		
source	1..760		
	/organism="Mus musculus molossinus"		
	/mol_type="genomic DNA"		
	/sub_species="molossinus"		
	/db_xref="taxon:57486"		
	/clone="MSMg01-264D02.TJ"		
	/sex="male"		
	/tissue_type="mixture of kidney and spleen"		
	/clone_lib="MSMg01 Mouse Male BAC Library"		
ORIGIN			
Query Match	87.2%;	Score 21.8;	DB 10; Length 760;
Best Local Similarity	92.0%;	Pred. No. 32;	
Matches	23; Conservative	0; Mismatches	2; Indels 0; Gaps 0;
QY	1	AAAAAAGAGCGCAGACTGCTCTTCC	25
Db	576	AAAAAAGAGCGCAGCTGCTCTTCC	600
RESULT 5	BQ925815/c	938 bp	mRNA linear EST 20-AUG-2002
LOCUS	AGENCOURT 8801758 NIH_MGC_129		Mus musculus cDNA clone IMAGE:6315307
DEFINITION	5'', mRNA sequence.		
ACCESSION	BQ925815		


```

SOURCE
ORGANISM      Rattus norvegicus (Norway rat)
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE
AUTHORS      1 (bases 1 to 249)
TITLE        Bonaldo M.F., Lennon G. and Soares M.B.
JOURNAL      Normalization and subtraction: two approaches to facilitate gene
PUBMED      Genome Res. 6 (9), 791-806 (1996)
COMMENT      8889548
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics / NIH
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgaps-i@mail.nih.gov
Oligo-dT track not found, Not 1 site shown in beginning of sequence
is likely internal to the message. cDNA library Preparation: M.B.
Soares Lab Clone distribution: Clones will be available through
IMAGE (http://image.llnl.gov)
Seq primer: M13 FORWARD
POLYA=Yes.
FEATURES
source
1..249
Location/Qualifiers
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="IMAGE:7361685"
/dev_stage="ADULT"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI CGAP FS1"
/notes="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not 1; Site 2: Eco RI; UI-R-FS1 is a
normalized cDNA library constructed from Swam Rat
Chondrosarcoma. The library was constructed and
normalized according to Bonaldo, Lennon and Soares, Genome
Research, 6:791-806, 1996. The Rat SRC-JWS Cell Line tumor
was provided by Dr Jeff Stevens of the University of Iowa.
TAG TISSUE=Rat SRC-JWS Cell Line
TAG LIB=UI-R-FS1
TAG_SEQ=AGCGCCGAT"
ORIGIN
Query Match      80.8%; Score 20.2; DB 6; Length 249;
Best Local Similarity 88.0%; Pred. No. 1.6e+02;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY      1 AAAAAAGCGCAGACTGGCTCTCC 25
Db      173 AAAAAAGAGCAGTACTGCTCTCC 149
RESULT 8
BZ192915/c
LOCUS      BZ192915      330 bp      DNA      linear      GSS 11-OCT-2002
DEFINITION CH230-248022.TJ CHORI-230 Segment 2 Rattus norvegicus genomic clone
CH230-248022, genomic survey sequence.
ACCESSION      BZ192915
VERSION      BZ192915.1 GI:23850972
KEYWORDS      GSS.
SOURCE      Rattus norvegicus (Norway rat)
ORGANISM      Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE
AUTHORS      1 (bases 1 to 330)
Zhao S., Shetty J., Shatsman S., Teegave G., Geer K.,
Shivatsbeyn A., Gebregorgis E., Overton L., Russell D., Chen D.,
Riggs F., de Jong P. and Fraser C.M.
Rat BAC End Sequences from Library CHORI-230 MboI segment
Unpublished (1999)
Other_GSSs: CH230-248022.TV
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the rat BAC library CHORI-230
(http://www.chori.org/bacpac/rat230.htm). For BAC library
availability, please contact Pieter de Jong (pdejong@mail.cho.org).
Clones may be purchased from BACPAC Resources
(http://www.chori.org/bacpac/or ering information.htm). BAC end
page: http://www.tigr.org/tdb/bac_ends/rat/bac_end_intro.html
Plate: 248 row: O column: 22
Seq primer: SP6
Class: BAC ends.
FEATURES
Location/Qualifiers
1..330
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/strain="BN/SsNHsd/MCW"
/db_xref="taxon:10116"
/clone="CH230-248022"
/sex="Female"
/cell_type="Brain"
/clone_lib="CHORI-230 Segment 2"
/notes="Vector: pTARBAC1-3; Site 1: MboI; Site 2: MboI;
CHORI-230 Rat (BN/SsNHsd/MCW) BAC library produced by
Pieter de Jong"
ORIGIN
Query Match      80.8%; Score 20.2; DB 9; Length 330;
Best Local Similarity 88.0%; Pred. No. 1.6e+02;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY      1 AAAAAAGCGCAGACTGGCTCTCC 25
Db      258 AAAAAAGAGCCCTGACTGCTCTCC 234
RESULT 9
AU018134/c
LOCUS      AU018134      347 bp      mRNA      linear      EST 19-OCT-1998
DEFINITION AU018134 Mouse two-cell stage embryo cDNA Mus musculus cDNA clone
J0749B12 3', mRNA sequence.
ACCESSION      AU018134
VERSION      AU018134.1 GI:3373624
KEYWORDS      EST.
SOURCE      Mus musculus (house mouse)
ORGANISM      Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS      1 (bases 1 to 347)
Ko, M.S.H., Kitchen, J.R., Wang, X., Threat, T.A., Sun, T., DePalma, G.E.,
Liang, Y., Kargul, G.J., Sharara, R. and Doi, H.
Systematic analyses of genes expressed in 2-cell stage mouse
embryos (The ERATO/Doi Project at Wayne State University)
(Ko, M.S.H. et al.)
Unpublished (1998)
JOURNAL      Contact: Hirofumi Doi
COMMENT      Doi Biosymmetry Project, ERATO
Japan Science and Technology Corporation (JST)
WBG Marive East 12F, 2-6 Nakase, Mihama-ku, Chiba 261-71, Japan
Email: hdo@boa.jst.go.jp.
Location/Qualifiers
1..347
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="J0749B12"
/dev_stage="two-cell stage embryo"

```

/clone_lib="Mouse two-cell stage embryo cDNA"

ORIGIN

Query Match 80.8%; Score 20.2; DB 1; Length 347;
 Best Local Similarity 88.0%; Pred. No. 1.6e+02;
 Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 AAAAAAGAGCGCAGACTGGTCTTCC 25
 ||||| ||||| ||||| ||||| |||||
 Db 168 AAAAAAGAGCGCAGCTGCTCTTCC 144

RESULT 10

AZ947222/c 352 bp DNA linear GSS 27-APR-2001
 LOCUS 2M0209B21F Mouse 10kb plasmid UGCG2M library Mus musculus genomic
 DEFINITION clone UGCG2M0209B21 F, genomic survey sequence.

ACCESSION

AZ947222

VERSION

AZ947222.1 GI:13817987

KEYWORDS

GSS.

SOURCE

Mus musculus (house mouse)

ORGANISM

Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muridea; Muridae; Murinae; Mus.
 1 (bases 1 to 352)
 Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,
 Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T.,
 Reilly, M., Rose, R., Stokes, R., Tingey, A., von
 Niederhausern, A., and Wright, D., Weiss, R.

TITLE

Mouse whole genome scaffolding with paired end reads from 10kb

JOURNAL

Plasmid inserts

COMMENT

Unpublished (2000)
 Contact: Robert B. Weiss
 University of Utah Genome Center
 University of Utah
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0209 row: B column: 21
 Seq primer: CGTGTAAACGACGCGCAGT
 Class: plasmid ends
 High quality sequence stop: 352.

FEATURES

source

1. .352
 Location/Qualifiers
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UGCG2M0209B21"
 /sex="Female"

/lab host="E. coli strain XL10-Gold, T1-resistant, F-"
 /clone_lib="Mouse 10kb plasmid UGCG2M library"
 /note="Vector: PWD42nv; Purified genomic DNA from M.
 musculus C57BL/6J (female) was obtained from the Jackson
 Laboratory Mouse DNA Resource
 (http://www.jax.org/resources/documents/dnares/). The DNA
 was hydrodynamically sheared by repeated passage through a
 0.005 inch orifice at constant velocity. The sheared DNA
 was blunt end-repaired with T4 DNA polymerase and T4
 polynucleotide kinase. Adaptor oligonucleotides were
 ligated to the blunt ends in high molar excess. The
 adaptor DNA was purified and size-selected for a 9.5 to
 10.5 kb range using preparative agarose gel
 electrophoresis. Vector DNA was prepared from a derivative
 of PWD42 [gi4732114|gb|AF129072.1], a copy-number
 inducible derivative of plasmid R1. The vector was ligated
 with adaptors complementary to the insert adaptors and
 purified. The sheared, adaptor mouse DNA was annealed to
 adaptor vector DNA, and transformed into

chemically-competent E. coli XL10-Gold (Stratagene) cells
 and selected for ampicillin resistance."

ORIGIN

Query Match 80.8%; Score 20.2; DB 9; Length 352;
 Best Local Similarity 88.0%; Pred. No. 1.6e+02;
 Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 AAAAAAGAGCGCAGACTGGTCTTCC 25
 ||||| ||||| ||||| ||||| |||||
 Db 81 AAAAAAGAGCGCAGCTGCTCTTCC 57

RESULT 11

BG071090/c
 LOCUS BG071090.1

DEFINITION

H3094F01-3 NIA Mouse 15K cDNA Clone Set Mus musculus cDNA clone

ACCESSION

H3094F01.3

VERSION

BG071090

KEYWORDS

EST.

SOURCE

Mus musculus (house mouse)

ORGANISM

Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muridea; Muridae; Murinae; Mus.
 1 (bases 1 to 358)
 Kargul, G.J., Dudekula, D.B., Qian, Y., Lim, M.K., Jaradat, S.A.,
 Tanaka, T.S., Carter, M.G. and Ko, M.S.H.
 Verification and initial annotation of NIA mouse 15K cDNA clone set
 Unpublished (2001)

REFERENCE

Other ESTs: H3094F01-5

AUTHORS

Contact: George J. Kargul

TITLE

Laboratory of Genetics

JOURNAL

National Institute on Aging/National Institutes of Health

COMMENT

333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
 Email: cdna@lgsun.grc.nia.nih.gov
 This clone set has been freely distributed to the community. Please
 visit <http://lgsun.grc.nia.nih.gov/cDNA/15k.html> for details.
 Plate: H3094 row: F column: 01
 Seq primer: -21M13 Forward
 High quality sequence stop: 358
 POLYA=Yes.

FEATURES

source

1. .358
 Location/Qualifiers
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="niaEST:H3094F01-3"
 /db_xref="taxon:10090"
 /clone="H3094F01"
 /sex="Clones arrayed from a variety of cDNA libraries"
 /dev_stage="Clones arrayed from a variety of cDNA libraries"
 /lab host="DH10B"
 /clone_lib="NIA Mouse 15K cDNA Clone Set"
 /note="Vector: pSPORT1; Site 1: Sali; Site 2: NotI; This
 clone is among a rearranged set of 15,247 clones from 11
 embryo cDNA libraries (including preimplantation stage
 embryos from unfertilized egg to blastocyst, embryonic
 part of E7.5 embryos, extraembryonic part of E7.5
 embryos, and E12.5 female mesonephros/gonad) and one
 newborn ovary cDNA library. Average insert size 1.5 kb.
 All source libraries are cloned unidirectionally with
 Oligo(dT)-Not primers. References include: (1)
 Genome-wide expression profiling of mid-gestation
 placenta and embryo using a 15,000 mouse developmental
 cDNA microarray. 2000, Proc. Natl. Acad. Sci. U S A, 97:
 9127-9132; (2) Large-scale cDNA analysis reveals phased
 gene expression patterns during preimplantation mouse
 development, 2000, Development, 127: 1737-1749; (3)
 Genome-wide mapping of unselected transcripts from
 extraembryonic tissue of 7.5-day mouse embryos reveals

enrichment in the t-complex and under-representation on the X chromosome, 1998, Hum Mol Genet 7: 1967-1978."

ORIGIN
Query Match 80.8%; Score 20.2; DB 2; Length 358;
Best Local Similarity 88.0%; Pred. No. 1.6e+02;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 AAAAAAGAGCGAGCTGCTCTCC 25
|||||
Db 168 AAAAAAGAGCGAGCTGCTCTCC 144
|||||

RESULT 12
BH315230/c
LOCUS
DEFINITION
CH230-159C13-TV CHORI-230 Segment 1 Rattus norvegicus genomic clone
CH230-159C13, genomic survey sequence.
ACCESSION
BH315230
VERSION
BH315230.1 GI:17245926
KEYWORDS
GSS.
SOURCE
Rattus norvegicus (Norway rat)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidae; Muridae; Murinae; Rattus.

REFERENCE
AUTHORS
Zhao, S., Shetty, J., Shatsman, S., Teegaye, G., Geer, K.,
Shvartsbeyn, A., Gabregorgis, E., Overton, L., Russell, D., Chen, D.,
Riggs, F., de Jong, P. and Fraser, C.M.
Rat BAC End Sequences from Library CHORI-230 EcoRI segment
Unpublished (1999)
Other GSSs: CH230-159C13.TJ

TITLE
JOURNAL
COMMENT
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the rat BAC library CHORI-230
(http://www.chori.org/bacpac/rat230.htm). For BAC library
availability, please contact Pieter de Jong (pdejong@email.cho.org).
Clones may be purchased from BACPAC Resources
(http://www.chori.org/bacpac/or ering_information.htm). BAC end
page: http://www.tigr.org/tdb/bac_ends/rat/bac_end_intro.html
Plate: 159 row: C column: 13
Seq primer: T7
Class: BAC ends.

FEATURES
source
Location/Qualifiers
1..363
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/strain="BN/SENHed/MCW"
/db_xref="taxon:10116"
/clone="CH230-159C13"
/sex="Female"
/cell_type="Brain"
/clone_lib="CHORI-230 Segment 1"
/notes="Vector: PTARBAC2.1; Site 1: EcoRI; Site 2: EcoRI;
CHORI-230 Rat (BN/SENHed/MCW) BAC library produced by
Pieter de Jong"

ORIGIN
Query Match 80.8%; Score 20.2; DB 9; Length 363;
Best Local Similarity 88.0%; Pred. No. 1.6e+02;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 AAAAAAGAGCGAGCTGCTCTCC 25
|||||
Db 322 AAAAAAGAGCGAGCTGCTCTCC 298
|||||

RESULT 13
BG083952
LOCUS
DEFINITION
H3094F01-5 NIA Mouse 15K cDNA Clone Set Mus musculus cDNA clone
H3094F01 5', mRNA sequence.
ACCESSION
BG083952
VERSION
BG083952.1 GI:12566571
KEYWORDS
EST.
SOURCE
Mus musculus (house mouse)

ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidae; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 371)

AUTHORS
Kargul, G.J., Dudekula, D.B., Qian, Y., Lim, M.K., Jaradat, S.A.,
Tanaka, T.S., Carter, M.G. and Ko, M.S.H.
Verification and initial annotation of NIA mouse 15K cDNA clone set
Unpublished (2001)
Other ESTs: H3094F01-3
Contact: George J. Kargul
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Casell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@lgsun.grc.nia.nih.gov
This clone set has been freely distributed to the community. Please
visit http://lgsun.grc.nia.nih.gov/cDNA/15k.html for details.
Plate: H3094 row: F column: 01
Seq primer: -21M13 Reverse
High quality sequence stop: 371
POLYA=No.

FEATURES
source
Location/Qualifiers
1..371
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="niaEST:H3094F01-5"
/db_xref="taxon:10090"
/clone="H3094F01"
/sex="Clones arrayed from a variety of cDNA libraries"
/dev_stage="Clones arrayed from a variety of cDNA libraries"
/lab_host="DH10B"
/clone_lib="NIA Mouse 15K cDNA Clone Set"
/notes="Vector: pSPORT1; Site 1: SalI; Site 2: NotI; This
clone is among a rearrayed set of 15,247 clones from 11
embryo cDNA libraries (including preimplantation stage
embryos from unfertilized egg to blastocyst, embryonic
part of E7.5 embryos, extraembryonic part of E7.5
embryos, and E12.5 female mesonephros/gonad) and one
newborn ovary cDNA library. Average insert size 1.5 kb.
All source libraries are cloned unidirectionally with
Oligo(dT)-Not primers. References include: (1)
Genome-wide expression profiling of mid-gestation
placenta and embryo using a 15,000 mouse developmental
cDNA microarray, 2000, Proc. Natl. Acad. Sci. U S A, 97:
9127-9132; (2) Large-scale cDNA analysis reveals phased
gene expression patterns during preimplantation mouse
development, 2000, Development, 127: 1737-1749; (3)
Genome-wide mapping of unselected transcripts from
extraembryonic tissue of 7.5-day mouse embryos reveals
enrichment in the t-complex and under-representation on
the X chromosome, 1998, Hum Mol Genet 7: 1967-1978."

ORIGIN
Query Match 80.8%; Score 20.2; DB 2; Length 371;
Best Local Similarity 88.0%; Pred. No. 1.6e+02;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 AAAAAAGCGAGCTGCTCTCC 25
|||||
Db 180 AAAAAAGCGAGCTGCTCTCC 204
|||||

RESULT 14
BH344887/c
LOCUS
DEFINITION CH230-183013.TJ CHORI-230 Segment 1 Rattus norvegicus genomic clone
CH230-183013, genomic survey sequence.
ACCESSION BH344887
VERSION BH344887.1 GI:17275621
KEYWORDS
SOURCE Rattus norvegicus (Norway rat)
ORGANISM
REFERENCE
AUTHORS Zhao S., Shetty J., Shatsman S., Tsegaye G., Geer K., Shvartsbeyn A., Gebregorgis B., Overton L., Russell D., Chen D., Riggs F., de Jong P. and Fraser C.M.
TITLE Rat BAC End Sequences from Library CHORI-230 EcoRI segment
JOURNAL Unpublished (1999)
COMMENT Other GSSs: CH230-183013.TV
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: shao@tigr.org
Clones are derived from the rat BAC library CHORI-230
(http://www.chori.org/bacpac/rat230.htm). For BAC library
availability, please contact Pieter de Jong (pdejong@mail.cho.org).
Clones may be purchased from BACPAC Resources
(http://www.chori.org/bacpac/or ering information.htm). BAC end
page: http://www.tigr.org/tdb/bac_ends/rat/bac_end_intro.html
Plate: 183 row: 0 column: 13
Seq primer: SP6
Class: BAC ends.

FEATURES
source
1..393
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/strain="BN/SENHad/MCW"
/db_xref="taxon:10116"
/clones="CH230-183013"
/sex="Female"
/cell_type="Brain"
/clone_lib="CHORI-230 Segment 1"
/note="Vector: pTARBAC2.1; Site 1: EcoRI; Site 2: EcoRI;
CHORI-230 Rat (BN/SENHad/MCW) BAC library produced by
Pieter de Jong"

ORIGIN
Query Match 80.8%; Score 20.2; DB 9; Length 393;
Best Local Similarity 88.0%; Pred. No. 1.7e+02;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AAAAAAGCGCAGACTGGTCTTCC 25
|||||
Db 332 AAAAAAGAGCAGCCGACTGCTCTTCC 308
|||||

RESULT 15
CN669196
LOCUS
DEFINITION CN669196 NTA Mouse E13.5 whole embryo cDNA library (Long) Mus
musculus cDNA clone NIA:A0876H02 IMAGE:30764533 5', mRNA sequence.
ACCESSION CN669196
VERSION CN669196.1 GI:47435647
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.

REFERENCE
AUTHORS

1 (bases 1 to 400)
Sharov A.A., Piao Y., Matoba R., Dudekula D.B., Qian Y.,
VanBuren V., Falco G., Martin P.R., Stagg C.A., Basse Y.,
Wang Y., Carter M.G., Hamatani T., Aiba K., Akutsu H., Sharova L.,
Tanaka T.S., Kimber W.L., Yoshikawa T., Jaradat S.A., Pantano S.,
Nagaraja R., Boheler K.R., Raub D., Hodes R.J., Longo D.L.,
Schlessinger D., Keller J., Klotz E., Kelsoe G., Umazawa A.,
Vescovi A.L., Rossant J., Kunath T., Hogan B.L., Curcio A.,
D'Urso M., Kelsoe J., Hide W. and Ko M.S.

TITLE
JOURNAL
COMMENT

Transcriptome analysis of mouse stem cells and early embryos
PLoS Biol. 1 (3), 410-419 (2003)
Contact: Dawood B. Dudekula
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@nigun.grc.nia.nih.gov
Plate: A0876 row: H column: 02
Seq primer: M13 Reverse
High quality sequence stop: 400
POLYA=No.

FEATURES
source

1..400
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="niaEST:A0876H02-5"
/db_xref="taxon:10090"
/clone="NIA:A0876H02 IMAGE:30764533"
/tissue_type="whole embryo including extraembryonic
tissues at 13.5-days postcoitum"
/dev_stage="E13.5"
/lab_host="DH10B"
/clone_lib="NIA Mouse E13.5 whole embryo cDNA library
(Long)"
/note="Vector: pCMV-SPORT6 (Invitrogen); Site 1: SalI;
Site 2: NotI; Mouse cDNA project by the Laboratory of
Genetics, National Institute on Aging (NIA), Intramural
Research Program, NIH (http://lgsun.grc.nia.nih.gov/cDNA).
This is a long-transcript enriched cDNA library (Ref.
Genome Res. 11: 1553-1558 (2001). [PMID: 11544199]). Total
RNAs were extracted from 1 embryo at 13.5-days postcoitum.
Double-stranded cDNAs were synthesized with an Oligo(dT)
primer [Invitrogen]:
5'-pGACTAGTTCAGATCGCGCGCCCTTTT-3' from
3ug of total RNA, treated with T4 DNA polymerase, and
purified by ethanol-precipitation. The cDNAs were ligated
to lone-linker L1-Sal4, purified by phenol/chloroform, and
separated from free linkers by Centricon 100. Then, the
cDNAs were amplified by long-range high fidelity PCR using
Ex Taq polymerase (Takara) with a primer Sal4-S. The
products were purified by phenol/chloroform and Centricon
100. The cDNAs were digested with SalI and NotI enzymes
and cloned into SalI/NotI site of pCMV-SPORT6 plasmid
vector. The DH10B E. coli host was transformed with the
ligation mixture by the standard chemical method. The
average insert size is about 3.0Kb. The library was
constructed by Yulan Piao."

ORIGIN

Query Match 80.8%; Score 20.2; DB 7; Length 400;
Best Local Similarity 88.0%; Pred. No. 1.7e+02;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AAAAAAGCGCAGACTGGTCTTCC 25
|||||
Db 30 AAAAAAGAGCAGCTGCTCTTCC 54
|||||

RESULT 16
LOCUS

AW535711/c
DEFINITION UI-R-BS0-aod-b-06-0-UI.sl UI-R-BS0 Rattus norvegicus cDNA clone
417 bp mRNA linear EST 06-MAR-2000
UI-R-BS0-aod-b-06-0-UI 3', mRNA sequence.

Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A., Kuroji, H., Kawasawa, Y., Kedzierski, R.M., King, B.D., Konagaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W.J., Perte, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S., Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M., Sandelin, A., Schneider, C., Semple, C.A., Setou, M., Shimada, M., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M., Verardo, R., Wagner, L., Wahlested, C., Wang, Y., Watanabe, Y., Wells, C., Wu, Z., Wynshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S., Rogers, J., Birney, E. and Hayashizaki, Y.

Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
12456851

TITLE
JOURNAL
PUBLISHED
COMMENT

Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216

Query Match	80.8%	Score	20.2;	DB	5;	Length	430;
Best Local Similarity	88.0%;	Pred. No.	1.7e+02;				
Matches	22;	Conservative	0;	Mismatches	3;	Indels	0;
						Gaps	0;

Qy	RESULT 19	REFERENCE
Db	BP417531/c	AUTHORS
	LOCUS	TITLE
	DEFINITION	JOURNAL
	ACCESSION	COMMENT
	VERSION	
	KEYWORDS	
	SOURCE	
	ORGANISM	

ORIGIN

Qy Db

RESULT 20	ACCESSION
AZ960237/c	VERSION
LOCUS	KEYWORDS
DEFINITION	SOURCE
	ORGANISM

[illegible]

84112, USA
Tel: 801 585 5606
Fax: 801 585 7177

Email: dunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0228 row: P column: 09

Seq primer: CGTTGTAACGACGCGCCAGT

Class: plasmid ends

High quality sequence stop: 464.

FEATURES

source

1. .464

/organism="Mus musculus"

/mol_type="genomic DNA"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="UUGC2M0228P09"

/sex="Female"

/lab_host="E. coli strain XL10-Gold, TI-resistant, F-"

/clone_lib="Mouse 10kb plasmid UUGC2M library"

/notes="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (female) was obtained from the Jackson Laboratory Mouse DNA Resource

(http://www.jax.org/resources/documents/dnares/). The DNA

was hydrodynamically sheared by repeated passage through a

0.005 inch orifice at constant velocity. The sheared DNA

was blunt end-repaired with T4 DNA polymerase and T4

polynucleotide kinase. Adaptor oligonucleotides were

ligated to the blunt ends in high molar excess. The

adapted DNA was purified and size-selected for a 9.5 to

10.5 kb range using preparative agarose gel

electrophoresis. Vector DNA was prepared from a derivative

of pWD42 (G1[4732114]9b|AF129072.1), a copy-number

inducible derivative of plasmid R1. The vector was ligated

with adaptors complementary to the insert adaptors and

purified. The sheared, adapted mouse DNA was annealed to

adapted vector DNA, and transformed into

chemically-competent E. coli XL10-Gold (Stratagene) cells

and selected for ampicillin resistance."

ORIGIN

Query Match

Best Local Similarity

Matches

22; Conservative

0; Mismatches

3; Indels

0; Gaps

0;

QY 1 AAAAAAGCGCAGACTGGTCTTCC 25

Db 413 AAAAAAGCGCAGCTGACTGCTCTTCC 389

RESULT 21

BH260028/c

LOCUS

DEFINITION

CH230-113A8-TV CHORI-230 Segment 1 Rattus norvegicus genomic clone

CH230-113A8, genomic survey sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

Sciurognathi; Muridea; Muridae; Murinae; Rattus.

1 (bases 1 to 480)

Shvartsbeyn, A., Gebregeorgis, E., Overton, L., Russell, D., Chen, D.,

Riggs, F., de Jong, P. and Fraser, C.M.

Rat BAC End Sequences from Library CHORI-230 EcoRI segment

Unpublished (1999)

Other GSSs: CH230-113A8.TJD

Contact: Shaying Zhao

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208

Email: szhao@igr.org

Clones are derived from the rat BAC library CHORI-230

(http://www.chori.org/bacpac/rat230.htm). For BAC library

availability, please contact Pieter de Jong (pdejong@mail.cho.org).

Clones may be purchased from BACPAC Resources

(http://www.chori.org/bacpac/or ering information.htm). BAC end

page: http://www.tigr.org/tdb/bac_ends/rat/bac_end_intro.html

Plate: 113 row: A column: 8

Seq primer: T7

Class: BAC ends.

Location/Qualifiers

1. 480

/organism="Rattus norvegicus"

/mol_type="genomic DNA"

/strain="BN/SSNHsd/MCW"

/db_xref="taxon:10116"

/clone="CH230-113A8"

/sex="Female"

/cell_type="Brain"

/clone_lib="CHORI-230 Segment 1"

/note="Vector: pTARBAC2.1; Site 1: EcoRI; Site 2: EcoRI;

CHORI-230 Rat (BN/SSNHsd/MCW) BAC library produced by

Pieter de Jong"

ORIGIN

Query Match

Best Local Similarity

Matches

22; Conservative

0; Mismatches

3; Indels

0; Gaps

0;

QY 1 AAAAAAGCGCAGACTGGTCTTCC 25

Db 174 AAAAAAGCGCAGCTGCTCTTCC 150

RESULT 22

BY569957/c

LOCUS

DEFINITION

CDNA clone F830216K22 3', mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Mus musculus

(house mouse)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

Sciurognathi; Muroidea; Muridae; Murinae; Mus.

1 (bases 1 to 481)

Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,

Nikaido, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I.,

Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,

Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C.,

Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H.,

Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusic, V.,

Chothia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A.,

Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T.,

Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S.,

Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A.,

Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R.,

Maltais, L., Marchionni, L., Mckenzie, L., Miki, H., Nagashima, T.,

Numata, K., Okido, T., Pavan, W.J., Pertea, G., Pesole, G.,

Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S.,

Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M.,

Sandelin, A., Schneider, C., Semple, C.A., Setou, M., Shimada, K.,

Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M.,

Verardo, R., Wagner, L., Wahlestedt, C., Wang, X., Watanabe, Y.,

Wells, C., Wilming, L.G., Wysshaw-Boris, A., Yanagisawa, M., Yang, I.,

Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P.,

Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M.,

Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,

BY569957 481 bp mRNA linear EST 15-DEC-2002

BY569957 RIKEN full-length enriched, activated spleen Mus musculus

CDNA clone F830216K22 3', mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Mus musculus

(house mouse)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

Sciurognathi; Muroidea; Muridae; Murinae; Mus.

1 (bases 1 to 481)

Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,

Nikaido, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I.,

Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,

Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C.,

Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H.,

Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusic, V.,

Chothia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A.,

Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T.,

Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S.,

Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A.,

Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R.,

Maltais, L., Marchionni, L., Mckenzie, L., Miki, H., Nagashima, T.,

Numata, K., Okido, T., Pavan, W.J., Pertea, G., Pesole, G.,

Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S.,

Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M.,

Sandelin, A., Schneider, C., Semple, C.A., Setou, M., Shimada, K.,

Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M.,

Verardo, R., Wagner, L., Wahlestedt, C., Wang, X., Watanabe, Y.,

Wells, C., Wilming, L.G., Wysshaw-Boris, A., Yanagisawa, M., Yang, I.,

Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P.,

Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M.,

Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,

non-normalized rat tongue library cDNA Library Preparation: M.B. Soares Lab Clone distribution: clones will be available through Research Genetics (www.resgen.com) The following repetitive elements were found in this cDNA sequence: 1-155, >B2#SINE/B2 164-255, >ID RN#SINE/ID 258-294, >B2#SINE/B2 299-493, >B2#SINE/B2 Seq primer: M13 Forward POLYA=Yes.

FEATURES

source
Location/Qualifiers
1. .506
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-DL0-cio-f-02-0-UI"
/dev_stage="ADULT"
/lab_host="DH10B (Life Technologies)"
/clone_lib="UI-R-DL0"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; The UI-R-DL0 library is a non-normalized Rat Tongue library constructed in pT7T3 PAC vector according to the procedure described by Bonaldo, Lennon & Soares (Genome Research 6: 791-806, 1996). The oligonucleotide used to prime first strand synthesis contained the sequence tag GCGAA between the Not I cloning site and dT18 stretch.
TAG_TISUE=rat tongue
TAG_LIB=UI-R-DL0
TAG_SEQ=GCGAA"

ORIGIN

Query Match 80.8%; Score 20.2; DB 3; Length 506;
Best Local Similarity 88.0%; Pred. No. 1.7e+02;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAAGAGCGCAGACTGGTCTTCC 25
|||||
DB 159 AAAAAAGAGCGCAGCTGCTCTTCC 135

RESULT 25
AZ875623/c
LOCUS
DEFINITION
2M0190D13F Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC2M0190D13 F, genomic survey sequence.
ACCESSION
AZ875623
VERSION
AZ875623.1 GI:13085816
KEYWORDS
GSS.
SOURCE
Mus musculus (house mouse)

ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
Unpublished (2000)

TITLE
Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL
Unpublished (2000)
COMMENT
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: dunn@genetics.utah.edu
Insert Length: 10000 Scd Error: 0.00
Plate: 0190 row: D column: 13
Seq primer: CGTTGTAACGACGGCCAGT
Class: plasmid ends
High quality sequence stop: 509.

FEATURES

source
Location/Qualifiers
1. .509
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M0190D13"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, Tl-resistant, P-"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

ORIGIN

Query Match 80.8%; Score 20.2; DB 9; Length 509;
Best Local Similarity 88.0%; Pred. No. 1.7e+02;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAAGAGCGCAGACTGGTCTTCC 25
|||||
DB 356 AAAAAAGAGCTGACTGCTCTTCC 332

RESULT 26
BB756913/c
LOCUS
DEFINITION
BB756913 RIKEN full-length enriched, melanocyte Mus musculus cDNA clone G270079J17 3', mRNA sequence.
ACCESSION
BB756913
VERSION
BB756913.1 GI:16203240
KEYWORDS
EST.
SOURCE
Mus musculus (house mouse)

ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 528)

AUTHORS
Akimura,T., Arakawa,T., Carninci,P., Furuno,M., Hanagaki,T., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Imotani,K., Ishii,Y., Ito,M., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Matuyama,T., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sakazume,N., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tonaru,A., Toya,T., Watahiki,A., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.
RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al. 2001)

TITLE
Unpublished (2001)
JOURNAL
COMMENT
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216

Email: genome-res@gsr.riken.jp, URL: <http://genome.gsc.riken.jp/>
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Sugahara, Y., Shibata, K.,
 Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new
 genes. Genome Res. 10 (10), 1617-1630 (2000)
 wagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,
 Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T.,
 Matsumura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A.
 and Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.
 10 (11), 1757-1771 (2000)
 Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P.,
 Sugahara, Y., and Hayashizaki, Y.
 Computer-based methods for the mouse full-length cDNA
 encyclopedia: real-time sequence clustering for construction of a
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 Please visit our web site (<http://genome.gsc.riken.go.jp>) for
 further details.

FEATURES source
 Location/Qualifiers
 1. 528
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="G270079J17"
 /cell_type="melanocyte"
 /clone_lib="RIKEN full-length enriched, melanocyte"
 Query Match 80.8%; Score 20.2; DB 2; Length 528;
 Best Local Similarity 88.0%; Pred. No. 1.7e+02;
 Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 Qy 1 AAAAAAGAGCGACTGCTCTCC 25
 |||||
 Db 383 AAAAAAGAGCGACTGCTCTCC 359
 |||||

RESULT 27
 BH349233/c
 LOCUS 536 bp DNA linear GSS 03-DEC-2001
 DEFINITION CH230-81B24.TJ CHORI-230 Segment 1 Rattus norvegicus genomic clone
 CH230-81B24, genomic survey sequence.
 ACCESSION BH349233
 VERSION BH349233.1 GI:17279967
 KEYWORDS GSS.
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muroidae; Muridae; Murinae; Rattus.
 1 (bases 1 to 536)
 REFERENCES
 AUTHORS Zhao, S., Shetty, J., Shatsman, S., Tsegaye, G., Geer, K.,
 Shwartsbeyn, A., Gebregorgis, E., Overton, L., Russell, D., Chen, D.,
 Riggs, P., de Jong, P., and Fraser, C.M.
 TITLE Rat BAC End Sequences from Library CHORI-230 EcoRI segment
 JOURNAL Unpublished (1999)
 COMMENT Other_GSSs: CH230-81B24.TV
 Contact: Shaying Zhao
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: szhao@tigr.org
 Clones are derived from the rat BAC library CHORI-230
 (<http://www.chori.org/bacpac/rat230.htm>). For BAC library
 availability, please contact Pieter de Jong (pdejong@mail.choi.org).
 Clones may be purchased from BACPAC Resources
 (http://www.chori.org/bacpac/orering_information.htm). BAC end

page: http://www.tigr.org/tdb/bac_ends/rat/bac_end_intro.html
 Plate: 81 row: B column: 24
 Seq primer: SP6
 Class: BAC ends.

FEATURES source

Location/Qualifiers
 1. 536
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /strain="BN/SSHsd/MCW"
 /db_xref="taxon:10116"
 /clone="CH230-81B24"
 /sex="Female"
 /cell_type="Brain"
 /clone_lib="CHORI-230 Segment 1"
 /notes="Vector: pTARBAC2.1; Site 1: EcoRI; Site 2: EcoRI;
 CHORI-230 Rat (BN/SSHsd/MCW) BAC library produced by
 Pieter de Jong"

ORIGIN

Query Match 80.8%; Score 20.2; DB 9; Length 536;
 Best Local Similarity 88.0%; Pred. No. 1.7e+02;
 Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 Qy 1 AAAAAAGAGCGACTGCTCTCC 25
 |||||
 Db 170 AAAAAAGAGCGACTGCTCTCC 146
 |||||

RESULT 28

BG088328
 LOCUS 574 bp mRNA linear EST 18-DEC-2003
 DEFINITION H3151G04-5 NIA Mouse 15K cDNA Clone Set Mus musculus cDNA clone
 H3151G04 5', mRNA sequence.

ACCESSION

VERSION BG088328.2 GI:40109258
 KEYWORDS EST.

SOURCE

Mus musculus (house mouse)

ORGANISM

Mus musculus
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muroidae; Muridae; Murinae; Mus.

REFERENCE

AUTHORS

Tanaka, T.S., Jaradat, S.A., Lim, M.K., Kargul, G.J., Wang, X.,
 Grahovac, M.J., Pantano, S., Sano, Y., Piao, Y., Nagaraja, R., Doi, H.,
 Wood, W.H. III, Becker, K.G., and Ko, M.S.H.

TITLE

Genome-wide expression profiling of mid-gestation placenta and
 embryo using a 15,000 mouse developmental cDNA microarray

JOURNAL

Proc. Natl. Acad. Sci. U.S.A. 97 (16), 9127-9132 (2000)

COMMENT

On Jan 26, 2001 this sequence version replaced gi:12570892.

Other_ESTs: H3151G04-3

Contact: George J. Kargul

Laboratory of Genetics

National Institute on Aging/National Institutes of Health

333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA

Email: cdna@igun.grc.nia.nih.gov

This clone set has been freely distributed to the community. Please

visit <http://igun.grc.nia.nih.gov/cdna/15k.html> for details.

Plate: H3151 row: G column: 04

Seq primer: -21M13 Reverse

High quality sequence stop: 574

POLYA-No.

Location/Qualifiers

1. 574

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="niaEST:H3151G04-5"

/clone="H3151G04"

/sex="Clones arrayed from a variety of cDNA libraries"

/dev_stage="Clones arrayed from a variety of cDNA libraries"

/lab_host="DH10B"
 /clone_lib="NIA Mouse 15K cDNA Clone Set"
 /note="Vector: pSPORT1; Site 1: SalI; Site 2: NotI; This clone is among a rearranged set of 15,247 clones from 11 embryo cDNA libraries (including preimplantation stage embryos from unfertilized egg to blastocyst, embryonic part of E7.5 embryos, extraembryonic part of E7.5 embryos, and E12.5 female mesonephros/gonad) and one newborn ovary cDNA library. Average insert size 1.5 kb. All source libraries are cloned unidirectionally with Oligo(dT)-Not primers. References include: (1) Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray. 2000. Proc. Natl. Acad. Sci. U S A. 97: 9127-9132; (2) Large-scale cDNA analysis reveals phased gene expression patterns during preimplantation mouse development, 2000, Development, 127: 1737-1749; (3) Genome-wide mapping of unselected transcripts from extraembryonic tissue of 7.5-day mouse embryos reveals enrichment in the t-complex and under-representation on the X chromosome, 1998, Hum Mol Genet 7: 1967-1978."

ORIGIN

Query Match 80.8%; Score 20.2; DB 2; Length 574;
 Best Local Similarity 88.0%; Pred. No. 1.8e+02;
 Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAAGAGCGCAGACTGCTCTTCC 25
 |||||
 Db 256 AAAAAAGAGCACACTACTGCTCTTCC 280

RESULT 29

AZ288000 577 bp DNA linear GSS 27-JUL-2000
 LOCUS RPCI-23-142E9.TV RPCI-23 Mus musculus genomic clone RPCI-23-142E9,
 DEFINITION genomic survey sequence.

ACCESSION AZ288000
 VERSION AZ288000.1 GI:9529786
 KEYWORDS GSS.
 SOURCE Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muridea; Muridae; Murinae; Mus.

REFERENCE

Akinret, B., Levins, M., McGann, S., Tsegaye, G., Geer, K., Krol, M., de
 Zhao, S., Niernan, W., Feldblyum, T., Malek, J., Shatsman, S.,
 1 (bases 1 to 577)
 Jong, P. and Fraser, C.M.

Mouse BAC End Sequences from Library RPCI-23

Unpublished (1999)

Other GSSs: RPCI-23-142E9.TJ

Contact: Shaying Zhao

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208

Email: szhao@tigr.org

Clones are derived from the mouse BAC library RPCI-23. For BAC
 library availability, please contact Pieter de Jong

(pieterdejong.med.buffalo.edu). Clones may be purchased from
 BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm)
 or from Resea ch Genetics (info@resgen.com). BAC end page:
 http://www.tigr.org/tdb/bac ends/mouse/bac_end_intro.html

Plate: 142 row: E column: 9

Seq primer: T7

Class: BAC ends.

Location/Qualifiers

1. .577

/organism="Mus musculus"

/mol_type="genomic DNA"

/strain="C57BL/6J"

FEATURES

source

/db_xref="taxon:10090"
 /clone="RPCI-23-142E9"
 /sex="Female"
 /lab_host="DH10B"
 /clone_lib="RPCI-23"
 /note="Organ: Kidney/Brain; Vector: pBACe3.6; Site 1:
 EcoRI; Site 2: EcoRI; Female C57BL/6J mouse kidney and/or
 brain genomic DNA was isolated and partially digested
 with a combination of EcoRI and EcoRI Methylase. Size
 selected DNA was cloned into the pBACe3.6 vector at the
 EcoRI sites. The ligation products were transformed into
 DH10B electrocompetent cells (BRL Life Technologies). "

ORIGIN

Query Match 80.8%; Score 20.2; DB 9; Length 577;

Best Local Similarity 88.0%; Pred. No. 1.8e+02;

Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAAGAGCGCAGACTGCTCTTCC 25

|||||

Db 455 AAAAAAGAGTGTGCTGACTGATCTTCC 479

RESULT 30

BH261566

LOCUS

DEFINITION BH261566 586 bp DNA linear GSS 30-NOV-2001
 CH230-90J19.TV CHORI-230 Segment 1 Rattus norvegicus genomic clone

ACCESSION BH261566

VERSION BH261566.1 GI:17167188

KEYWORDS GSS.

SOURCE Rattus norvegicus (Norway rat)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muridea; Muridae; Murinae; Rattus.

1 (bases 1 to 586)

Zhao, S., Shetty, J., Shatsman, S., Tsegaye, G., Geer, K.,
 Shvartsbeyn, A., Gebregorgis, E., Overton, L., Russell, D., Chen, D.,
 Riggs, F., de Jong, P. and Fraser, C.M.

Rat BAC End Sequences from Library CHORI-230 EcoRI segment

Unpublished (1999)

Other GSSs: CH230-90J19.TJ

Contact: Shaying Zhao

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208

Email: szhao@tigr.org

Clones are derived from the rat BAC library CHORI-230

(http://www.chori.org/bacpac/rat230.htm). For BAC library

availability, please contact Pieter de Jong (pdejong@mail.cho.org).

Clones may be purchased from BACPAC Resources

(http://www.chori.org/bacpac/or ering information.html). BAC end

page: http://www.tigr.org/tdb/bac ends/rat/bac_end_intro.html

Plate: 90 row: J column: 19

Seq primer: T7

Class: BAC ends.

Location/Qualifiers

1. .586

/organism="Rattus norvegicus"

/mol_type="genomic DNA"

/strain="BN/SSNHsd/MCW"

/db_xref="taxon:10116"

/clone="CH230-90J19"

/sex="Female"

/cell_type="Brain"

/clone_lib="CHORI-230 Segment 1"

/note="Vector: pTARBAC2.1; Site 1: EcoRI; Site 2: EcoRI;

CHORI-230 Rat (BN/SSNHsd/MCW) BAC library produced by

Pieter de Jong"

ORIGIN

Query Match 80.8%; Score 20.2; DB 9; Length 586;
 Best Local Similarity 88.0%; Pred. No. 1.8e+02;
 Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAAGAGCGCAGACTGCTCTTCC 25
 |||||
 Db 93 AAAAAAGAGCGCAGCTGCTCTTCC 117
 |||||

RESULT 31
 AZ357207/c 607 bp DNA linear GSS 02-OCT-2000
 LOCUS
 DEFINITION
 1M0098D08R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
 clone UUGC1M0098D08 R, genomic survey sequence.

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muroidea; Muridae; Murinae; Mus.

REFERENCE
 AUTHORS
 Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,
 Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T.,
 Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von
 Niederhausern, A., and Wright, D., Weises, R.

TITLE
 Mouse whole genome scaffolding with paired end reads from 10kb
 plasmid inserts

JOURNAL
 COMMENT
 Unpublished (2000)
 Contact: Robert B. Weiss
 University of Utah Genome Center
 University of Utah
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
 84112, USA

Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0098 row: D column: 08

Seq primer: CACACAGGAAACAGCTATGACC
 Class: plasmid ends

High quality sequence stop: 607.

Location/Qualifiers

1. .607
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUGC1M0098D08"
 /sex="Male"

/lab_host="E. Coli strain XL10-Gold, T1-resistant, P-"
 /clone_lib="Mouse 10kb plasmid UUGC1M library"
 /note="Vector: PWD42nv; Purified genomic DNA from M.
 musculus C57BL/6J (male) was obtained from the Jackson
 Laboratory Mouse DNA Resource
 (http://www.jax.org/resources/documents/dnares/). The DNA
 was hydrodynamically sheared by repeated passage through a
 0.005 inch orifice at constant velocity. The sheared DNA
 was blunt end-repaired with T4 DNA polymerase and T4
 polynucleotide kinase. Adaptor oligonucleotides were
 ligated to the blunt ends in high molar excess. The
 adaptor DNA was purified and size-selected for a 9.5 to
 10.5 kb range using preparative agarose gel
 electrophoresis. Vector DNA was prepared from a derivative
 of pWD42 [gi|4732114|gb|AF129072.1], a copy-number
 inducible derivative of plasmid R1. The vector was ligated
 with adaptors complementary to the insert adaptors and
 purified. The sheared, adaptor mouse DNA was annealed to
 adaptor vector DNA, and transformed into
 chemically-competent E. coli XL10-Gold (Stratagene) cells
 and selected for ampicillin resistance."

FEATURES
 source

ORIGIN

Query Match 80.8%; Score 20.2; DB 9; Length 607;
 Best Local Similarity 88.0%; Pred. No. 1.8e+02;
 Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAAGAGCGCAGACTGCTCTTCC 25
 |||||
 Db 289 AAAAAAGAGCGCAGCTGCTCTTCC 265
 |||||

RESULT 32
 CN542163/c

LOCUS
 DEFINITION
 UI-R-EA0-ckx-d-03-0-UI.s1 UI-R-EA0 Rattus norvegicus cDNA clone
 UI-R-EA0-ckx-d-03-0-UI 3', mRNA sequence.

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

Rattus norvegicus (Norway rat)
 Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muroidea; Muridae; Murinae; Rattus.

REFERENCE
 AUTHORS
 TITLE
 Normalization and subtraction: two approaches to facilitate gene
 discovery

Bonaldo, M.F., Lennon, G. and Soares, M.B.

Genome Res. 6 (9), 791-806 (1996)

JOURNAL

FURNED

COMMENT

Contact: Soares, MB
 Coordinated Laboratory for Computational Genomics
 University of Iowa
 375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565

Email: bento-soares@uiowa.edu

Tissue Procurement: Tom Freeman, Sanger Center

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Distribution information can be found at
 http://genome.uiowa.edu/distribution/rat.html
 The following repetitive elements were found
 in this cDNA
 sequence: 1-171, >L1PB2#LINE/L1 (matched complement) 165-281,
 >BC1.MM#scrna (matched complement) 282-321, >L1PB3#LINE/L1 (matched
 complement)

Seq primer: M13 FORWARD

POLYA=Yes.

Location/Qualifiers

1. .625
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 /clone="UI-R-EA0-ckx-d-03-0-UI"
 /dev_stage="adult"
 /lab_host="UI-R-EA0"
 /clone_lib="UI-R-EA0"
 /note="Vector: pT73D-Pac (Pharmacia) with a modified
 polylinker; Site 1: Not I; Site 2: Eco RI; UI-R-EA0 is a
 non-normalized Rat ileum library (RIL) constructed in
 pT73 PAC vector according to the procedure described by
 Bonaldo, Lennon & Soares (Genome Research Genome 6:
 791-806, 1996). The oligonucleotide used to prime first
 strand synthesis contained the sequence tag CGACACTCG
 between the Not I cloning site and dt18 stretch. The Rat
 ileum tissue was provided by Tom Freeman of the Sanger
 Center.

TAG TISSUE=ileum
 TAG LIB=UI-R-EA0
 TAG_SEQ=CGACACTCG"

ORIGIN

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
1 (bases 1 to 708)
Zhao,S., Shetty,J., Shatsman,S., Tsegaye,G., Gear,K.,
Shvartsbeyn,A., Gebregeorgis,E., Overton,L., Russell,D., Chen,D.,
Riggs,F., de Jong,P. and Fraser,C.M.
Rat BAC End Sequences from Library CHORI-230 MboI segment
Unpublished (1999)
Other GSSs: CH230-269H12.TJ
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the rat BAC library CHORI-230
(http://www.chori.org/bacpac/rat230.htm). For BAC library
availability, please contact Pieter de Jong (pdejong@mail.cho.org).
Clones may be purchased from BACPAC Resources
(http://www.chori.org/bacpac/or ering information.htm). BAC end
page: http://www.tigr.org/tdb/bac_ends/rat/bac_end_intro.html
Plate: 269 row: H column: 12
Seq primer: T7
Class: BAC ends.

FEATURES
source

Location/Qualifiers
1. .708
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/strain="BN/SsNHsd/MCW"
/db_xref="taxon:10116"
/clone="CH230-269H12"
/sex="Female"
/cell_type="Brain"
/clone_lib="CHORI-230 Segment 2"
/note="Vector: pTARAC1.3; Site:1: MboI; Site:2: MboI;
CHORI-230 Rat (BN/SsNHsd/MCW) BAC library produced by
Pieter de Jong"

ORIGIN

Query Match 80.8%; Score 20.2; DB 9; Length 708;
Best Local Similarity 88.0%; Pred. No. 1.8e+02;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 1 AAAAAAGCGCGCTGACTGCTTCC 25
|||||
Db 584 AAAAAAGCGCGCTGACTGCTTCC 560
|||||

RESULT 36
AG562735/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
PUBMED
AUTHORS
TITLE

AG562735 714 bp DNA linear GSS 23-DEC-2004
Mus musculus molossinus DNA, clone:MSMg01-482J18.TJ, genomic survey
sequence.
AG562735.1 GI:48323433
GSS.
Mus musculus molossinus (Japanese wild mouse)
Mus musculus molossinus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
1
Abe,K., Noguchi,H., Tagawa,K., Yuzuriha,M., Toyoda,A., Kojima,T.,
Ezawa,K., Saitou,N., Hattori,M., Sakaki,Y., Moriwaki,K. and
Shiroishi,T.
Contribution of Asian mouse subspecies Mus musculus molossinus to
genomic constitution of strain C57BL/6J, as defined by BAC-end
sequence-SNP analysis
Genome Res. 14 (12), 2439-2447 (2004)
15574823
2 (bases 1 to 714)
Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
Direct Submission

JOURNAL

COMMENT

Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa, 230-0045, Japan
(E-mail:hattori@gsc.riken.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
Clones are derived from the mouse BAC library MSMg01. For BAC
library availability, please contact Kunya Abe (abe@rtc.riken.jp).
Tsukuba Institute, Bio Resource Center,
The Institute of Physical and Chemical Research (RIKEN) 3-1-1
Koyadai, Tsukuba, 305-0074 Japan
phone: 81-298-36-9189, fax: 81-298-36-9199
e-mail: abe@rtc.riken.jp

PRIMERS

Sequencing : TJ

LIBRARY

Vector : pBACe3.6

R.Site 1 : BAC01

R.Site 2 : BAC01

Location/Qualifiers

1. .714
/organism="Mus musculus molossinus"
/mol_type="genomic DNA"
/sub_species="molossinus"
/db_xref="taxon:57486"
/clone="MSMg01-482J18.TJ"
/sex="male"
/tissue type="mixture of kidney and spleen"
/clone_lib="MSMg01 Mouse Male BAC Library"

ORIGIN

Query Match 80.8%; Score 20.2; DB 10; Length 714;
Best Local Similarity 88.0%; Pred. No. 1.8e+02;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 1 AAAAAAGCGCGCTGCTTCC 25
|||||
Db 519 AAAAAAGCGCGCTGCTTCC 495
|||||

RESULT 37

BG298680/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

Location/Qualifiers

1. .726

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

/clone="IMAGE:4511566"

High quality sequence stop: 726.

Plate: L1AM10395 row: c column: 23

High quality sequence stop: 726.

Location/Qualifiers

1. .726

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

/clone="IMAGE:4511566"

High quality sequence stop: 726.

Plate: L1AM10395 row: c column: 23

High quality sequence stop: 726.

Location/Qualifiers

1. .726

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

/clone="IMAGE:4511566"

High quality sequence stop: 726.

Plate: L1AM10395 row: c column: 23

High quality sequence stop: 726.

Location/Qualifiers

1. .726

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

/clone="IMAGE:4511566"

High quality sequence stop: 726.

Plate: L1AM10395 row: c column: 23

High quality sequence stop: 726.

Location/Qualifiers

1. .726

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

/clone="IMAGE:4511566"

High quality sequence stop: 726.

Plate: L1AM10395 row: c column: 23

High quality sequence stop: 726.

Location/Qualifiers

1. .726

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

/clone="IMAGE:4511566"

High quality sequence stop: 726.

Plate: L1AM10395 row: c column: 23

High quality sequence stop: 726.

Location/Qualifiers

1. .726

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

/clone="IMAGE:4511566"

High quality sequence stop: 726.

Plate: L1AM10395 row: c column: 23

High quality sequence stop: 726.

Location/Qualifiers

1. .726

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

/clone="IMAGE:4511566"

High quality sequence stop: 726.

Plate: L1AM10395 row: c column: 23

High quality sequence stop: 726.

Location/Qualifiers

1. .726

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

/clone="IMAGE:4511566"

High quality sequence stop: 726.

Plate: L1AM10395 row: c column: 23

High quality sequence stop: 726.

Location/Qualifiers

1. .726

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

/clone="IMAGE:4511566"

High quality sequence stop: 726.

Plate: L1AM10395 row: c column: 23

High quality sequence stop: 726.

Location/Qualifiers

1. .726

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

/clone="IMAGE:4511566"

High quality sequence stop: 726.

Plate: L1AM10395 row: c column: 23

High quality sequence stop: 726.

Location/Qualifiers

1. .726

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

/clone="IMAGE:4511566"

High quality sequence stop: 726.

Plate: L1AM10395 row: c column: 23

High quality sequence stop: 726.

Location/Qualifiers

1. .726

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

/clone="IMAGE:4511566"

High quality sequence stop: 726.

Plate: L1AM10395 row: c column: 23

High quality sequence stop: 726.

Location/Qualifiers

1. .726

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

/clone="IMAGE:4511566"

High quality sequence stop: 726.

Plate: L1AM10395 row: c column: 23

High quality sequence stop: 726.

Location/Qualifiers

1. .726

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

/clone="IMAGE:4511566"

High quality sequence stop: 726.

Plate: L1AM10395 row: c column: 23

High quality sequence stop: 726.

Location/Qualifiers

1. .726

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

/clone="IMAGE:4511566"

High quality sequence stop: 726.

Plate: L1AM10395 row: c column: 23

High quality sequence stop: 726.

Location/Qualifiers

1. .726

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

/clone="IMAGE:4511566"

High quality sequence stop: 726.

Plate: L1AM10395 row: c column: 23

High quality sequence stop: 726.

Location/Qualifiers

1. .726

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

/clone="IMAGE:4511566"

High quality sequence stop: 726.

Plate: L1AM10395 row: c column: 23

High quality sequence stop: 726.

Location/Qualifiers

1. .726

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

/clone="IMAGE:4511566"

High quality sequence stop: 726.

Plate: L1AM10395 row: c column: 23

High quality sequence stop: 726.

Location/Qualifiers

1. .726

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

/clone="IMAGE:4511566"

High quality sequence stop: 726.

Plate: L1AM10395 row: c column: 23

High quality sequence stop: 726.

Location/Qualifiers

1. .726

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

/clone="IMAGE:4511566"

High quality sequence stop: 726.

Plate: L1AM10395 row: c column: 23

High quality sequence stop: 726.

Location/Qualifiers

1. .726

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

/clone="IMAGE:4511566"

High quality sequence stop: 726.

Plate: L1AM10395 row: c column: 23

High quality sequence

/tissue type="retina"
 /lab host="DH10B (phage-resistant)"
 /clone lib="NIH_MGC_94"
 /note="Organ: eye; Vector: PCMV-SPORT6; Site_1: NotI;
 Site_2: SalI; Cloned unidirectionally; oligo-dt primed.
 Average insert size 3.3 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH_MGC Library."

ORIGIN

Query Match 80.8%; Score 20.2; DB 2; Length 726;
 Best Local Similarity 88.0%; Pred. No. 1.8e+02;
 Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAAGCGCAGACTGCTCTTCC 25
 |||||
 Db 167 AAAAAAGCAGCTGCTCTTCC 143

RESULT 38

AZ430769/c
 LOCUS 736 bp DNA linear GSS 03-OCT-2000
 DEFINITION 1M0215G22F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
 clone UUGC1M0215G22 F, genomic survey sequence.

ACCESSION AZ430769
 VERSION 1
 KEYWORDS GSS.

SOURCE Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muridea; Muridae; Murinae; Mus.

1 (bases 1 to 736)

Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
 Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,
 Reilly,M., Rose,R., Stokes,R., Tingey,A., von
 Niederhausern,A. and Wright,D. Weiss,R.

Mouse whole genome scaffolding with paired end reads from 10kb

Plasmid inserts

Unpublished (2000)

Contact: Robert B. Weiss

University of Utah Genome Center

University of Utah

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
 84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00

Plate: 0215 row: G column: 22

Seq primer: CGTGTAAACGACGGCCAGT

Class: plasmid ends

High quality sequence stop: 736.

Location/Qualifiers

FEATURES

source

1..736
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUGC1M0215G22"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
 /clone_lib="Mouse 10kb plasmid UUGC1M library"
 /note="Vector: PWD42nv; Purified genomic DNA from M.
 musculus C57BL/6J (male) was obtained from the Jackson
 Laboratory Mouse DNA Resource
 (http://www.jax.org/resources/documents/dnares/). The DNA
 was hydrodynamically sheared by repeated passage through a
 0.005 inch orifice at constant velocity. The sheared DNA
 was blunt end-repaired with T4 DNA polymerase and T4
 polynucleotide kinase. Adaptor oligonucleotides were
 ligated to the blunt ends in high molar excess. The
 adapted DNA was purified and size-selected for a 9.5 to

10.5 kb range using preparative agarose gel
 electrophoresis. Vector DNA was prepared from a derivative
 of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
 inducible derivative of plasmid R1. The vector was ligated
 with adaptors complementary to the insert adaptors and
 purified. The sheared, adapted mouse DNA was annealed to
 adaptor vector DNA, and transformed into
 chemically-competent E. coli XL10-Gold (Stratagene) cells
 and selected for ampicillin resistance."

ORIGIN

Query Match 80.8%; Score 20.2; DB 9; Length 736;
 Best Local Similarity 88.0%; Pred. No. 1.8e+02;
 Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAAGCGCAGACTGCTCTTCC 25
 |||||
 Db 343 AAAAAAGCAGCTGCTCTTCC 319

RESULT 39

BZ173903
 LOCUS 779 bp DNA linear GSS 11-OCT-2002
 DEFINITION CH230-493G21.TJ CHORI-230 Segment 2 Rattus norvegicus genomic clone
 CH230-493G21, genomic survey sequence.

ACCESSION BZ173903
 VERSION 1
 KEYWORDS GSS.

SOURCE Rattus norvegicus (Norway rat)

ORGANISM

Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muridea; Muridae; Murinae; Rattus.

1 (bases 1 to 779)

Zhao,S., Shetty,J., Shatsman,S., Tsengay,G., Geer,K.,
 Shvartsbeyn,A., Gebregeorgis,E., Overton,L., Russell,D.,
 Riggs,F., de Jong,P. and Fraser,C.M.

Rat BAC End Sequences from Library CHORI-230 MboI segment

Unpublished (1999)

Other_GSSs: CH230-493G21.TV

Contact: Shaying Zhao

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208

Email: szhao@tigr.org

Clones are derived from the rat BAC library CHORI-230

(http://www.chori.org/bacpac/rat230.htm). For BAC library

availability, please contact Pieter de Jong (pdejong@mail.cho.org).

Clones may be purchased from BACPAC Resources

(http://www.chori.org/bacpac/or ering_information.htm). BAC end

plate: http://www.tigr.org/tdb/bac_ends/rat/bac_end_intro.html

Plate: 493 row: G column: 21

Seq primer: SP6

Class: BAC ends

Location/Qualifiers

FEATURES

source

1..779
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /strain="BN/SSNHsd/MCW"
 /db_xref="taxon:10116"
 /clone="CH230-493G21"
 /sex="Female"
 /cell_type="Brain"
 /clone_lib="CHORI-230 Segment 2"
 /note="Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI;
 CHORI-230 Rat (BN/SSNHsd/MCW) BAC library produced by
 Pieter de Jong"

ORIGIN

Query Match 80.8%; Score 20.2; DB 9; Length 779;
 Best Local Similarity 88.0%; Pred. No. 1.8e+02;

Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AAAAAAGCGCAGACTGCTCTTCC 25
 |||||
 Db 324 AAAAAAGCGCAGACTGCTCTTCC 348
 |||||

RESULT 40
 BH292979
 LOCUS CH230-92B14.TJB CHORI-230 Segment 1 Rattus norvegicus genomic clone
 CH230-92B14, genomic survey sequence.

ACCESSION BH292979
 VERSION BH292979.1 GI:17205387
 KEYWORDS GSS.
 SOURCE Rattus norvegicus (Norway rat)

ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE 1 (bases 1 to 789)
 AUTHORS Zhao,S., Shetty,J., Shatsman,S., Teegaye,G., Geer,K.,
 Shvartsbeyn,A., Gebregeorgis,E., Overton,L., Russell,D., Chen,D.,
 Riggs,F., de Jong,P. and Fraser,C.M.

TITLE Rat BAC End Sequences from Library CHORI-230 EcoRI segment
 JOURNAL Rat BAC End Sequences from Library CHORI-230 EcoRI segment
 COMMENT Unpublished (1999)

Other_GSSs: CH230-92B14.TVB
 Contact: Shaying Zhao
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208

Email: szhao@tigr.org
 Clones are derived from the rat BAC library CHORI-230
 (http://www.chori.org/bacpac/rat230.htm). For BAC library
 availability, please contact Pieter de Jong (pdejong@mail.cho.org).
 Clones may be purchased from BACPAC Resources
 (http://www.chori.org/bacpac/or ering_information.htm). BAC end
 page: http://www.tigr.org/tdb/bac_ends/rat/bac_end_intro.html
 Plate: 92 row: B column: 14
 Seq primer: SP6
 Class: BAC ends.

FEATURES
 source
 1..789
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /strain="BN/SSHsd/MCW"
 /db_xref="taxon:10116"
 /clone="CH230-92B14"
 /sex="Female"
 /cell_type="Brain"
 /clone_lib="CHORI-230 Segment 1"
 /note="Vector: PTARBAC2.1; Site 1: EcoRI; Site 2: EcoRI;
 CHORI-230 Rat (BN/SSHsd/MCW) BAC library produced by
 Pieter de Jong"

ORIGIN
 Query Match 80.8%; Score 20.2; DB 9; Length 789;
 Best Local Similarity 88.0%; Pred. No. 1.8e+02;
 Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AAAAAAGCGCAGACTGCTCTTCC 25
 |||||
 Db 590 AAAAAAGCGCTGACTGCTCTTCC 614
 |||||

RESULT 41
 BZ277088
 LOCUS CH230-308G22.TJ CHORI-230 Segment 2 Rattus norvegicus genomic clone
 CH230-308G22, genomic survey sequence.

VERSION BZ277088.1 GI:24000638
 KEYWORDS GSS.
 SOURCE Rattus norvegicus (Norway rat)

ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE 1 (bases 1 to 809)
 AUTHORS Zhao,S., Shetty,J., Shatsman,S., Teegaye,G., Geer,K.,
 Shvartsbeyn,A., Gebregeorgis,E., Overton,L., Russell,D., Chen,D.,
 Riggs,F., de Jong,P. and Fraser,C.M.

TITLE Rat BAC End Sequences from Library CHORI-230 MboI segment
 JOURNAL Rat BAC End Sequences from Library CHORI-230 MboI segment
 COMMENT Unpublished (1999)

Other_GSSs: CH230-308G22.TV
 Contact: Shaying Zhao
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208

Email: szhao@tigr.org
 Clones are derived from the rat BAC library CHORI-230
 (http://www.chori.org/bacpac/rat230.htm). For BAC library
 availability, please contact Pieter de Jong (pdejong@mail.cho.org).

Clones may be purchased from BACPAC Resources
 (http://www.chori.org/bacpac/or ering_information.htm). BAC end
 page: http://www.tigr.org/tdb/bac_ends/rat/bac_end_intro.html
 Plate: 308 row: G column: 22
 Seq primer: SP6
 Class: BAC ends.

FEATURES
 source
 1..809
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /strain="BN/SSHsd/MCW"
 /db_xref="taxon:10116"
 /clone="CH230-308G22"
 /sex="Female"
 /cell_type="Brain"
 /clone_lib="CHORI-230 Segment 2"
 /note="Vector: PTARBAC1.3; Site 1: MboI; Site 2: MboI;
 CHORI-230 Rat (BN/SSHsd/MCW) BAC library produced by
 Pieter de Jong"

ORIGIN
 Query Match 80.8%; Score 20.2; DB 9; Length 809;
 Best Local Similarity 88.0%; Pred. No. 1.8e+02;
 Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AAAAAAGCGCAGACTGCTCTTCC 25
 |||||
 Db 583 AAAAAAGCGCTGCTCTTCC 607
 |||||

RESULT 42
 AG295893/c
 LOCUS AG295893
 DEFINITION Mus musculus molossinus (Japanese wild mouse) sequence.

ACCESSION AG295893
 VERSION AG295893.1 GI:47868847
 KEYWORDS GSS.
 SOURCE Mus musculus molossinus (Japanese wild mouse)

ORGANISM Mus musculus molossinus
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1
 AUTHORS Abe,K., Noguchi,H., Tagawa,K., Yuzuriha,M., Toyoda,A., Kojima,T.,
 Ezawa,K., Saitou,N., Hattori,M., Sakaki,Y., Moriaki,K. and
 Shiroishi,T.

TITLE Contribution of Asian mouse subspecies Mus musculus molossinus to
 genomic constitution of strain C57BL/6J, as defined by BAC-end

JOURNAL
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

sequence-SNP analysis
Genome Res. 14 (12), 2439-2447 (2004)
15574823
2 (bases 1 to 816)
Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
Direct Submission
Submitted (17-NOV-2003) Maehira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC), Japan
1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa, 230-0045,
(E-mail:hattori@gsc.riken.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
Clones are derived from the mouse BAC library MSMg01. For BAC
library availability, please contact Kuniya Abe (abe@rtc.riken.jp).
Tsukuba Institute, Bio Resource Center, Research (RIKEN) 3-1-1
The Institute of Physical and Chemical Research (RIKEN) 3-1-1
Koyadai, Tsukuba, 305-0074 Japan
phone: 81-298-36-9189, fax: 81-298-36-9199
e-mail: abe@rtc.riken.jp

PRIMERS
Sequencing : TJ
LIBRARY
Vector : pBACe3.6
R.Site 1 : EcoRI
R.Site 2 : EcoRI.

FEATURES
source
Location/Qualifiers
1..816
/organism="Mus musculus molossinus"
/mol_type="genomic DNA"
/sub_species="molossinus"
/db_xref="taxon:57486"
/clone="MSMg01-072F03.TJ"
/sex="male"
/tissue_type="mixture of kidney and spleen"
/clone_lib="MSMg01 Mouse Male BAC Library"

ORIGIN
Query Match 80.8%; Score 20.2; DB 10; Length 816;
Best Local Similarity 88.0%; Pred. No. 1.8e+02;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAAGAGCGAGACTGGTCTTCC 25
|||||
Db 351 AAAAAAGAGCGAGACTGGTGGCCC 327

RESULT 43
BQ959424/c
LOCUS
DEFINITION
AGENCY: BQ959424, 831 bp mRNA linear EST 21-AUG-2002
IMAGE: 6476660 5', mRNA sequence.
BQ959424
BQ959424.1 GI:22374902
EST.
Mus musculus (house mouse)

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Murinae; Mus.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
NITH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: The Cepko Laboratory
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM14017 Row: f column: 21
High quality sequence stop: 394.

FEATURES
source

Location/Qualifiers
1..831
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:6476660"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI CGAP Co24"
/note="Organ: colon; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.6 kb. Constructed by Life
Technologies. Note: this is a NCI CGAP Library."

ORIGIN

Query Match 80.8%; Score 20.2; DB 5; Length 831;
Best Local Similarity 88.0%; Pred. No. 1.8e+02;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAAGAGCGAGACTGGTCTTCC 25
|||||
Db 185 AAAAAAGAGCGAGACTGGTCTTCC 161

RESULT 44
B2275313
LOCUS
DEFINITION
CH230-450K22.TJ CHORI-230 Segment 2 Rattus norvegicus genomic clone
CH230-450K22, genomic survey sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Rattus norvegicus (Norway rat)

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Murinae; Rattus.
1 (bases 1 to 1052)
Zhao,S., Shetty,J., Shatsman,S., Tsegaye,G., Geer,K.,
Shvartsbeyn,A., Gebregeorgis,E., Overton,L., Russell,D., Chen,D.,
Riggs,F., de Jong,P. and Fraser,C.M.
Rat BAC End Sequences from Library CHORI-230 MboI segment
Unpublished (1999)
Other GSSs: CH230-450K22.TV
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org

Clones are derived from the rat BAC library CHORI-230
(http://www.chori.org/bacpac/rat230.htm). For BAC library
availability, please contact Pieter de Jong (pdejong@mail.cho.org).
Clones may be purchased from BACPAC Resources
(http://www.chori.org/bacpac/or ering information.htm). BAC end
page: http://www.tigr.org/cdb/bac_ends/rat/bac_end_intro.html
Plate: 450 row: K column: 22
Seq primer: SP6
Class: BAC ends.

FEATURES
source

Location/Qualifiers
1..1052
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/strain="BN/SSNHsd/MCW"
/db_xref="taxon:10116"
/clone="CH230-450K22"
/sex="Female"
/cell_type="Brain"
/clone_lib="CHORI-230 Segment 2"
/note="Vector: pTARBAC1.3; Site_1: MboI; Site_2: MboI;
CHORI-230 Rat (BN/SSNHsd/MCW) BAC library produced by
Pieter de Jong"

ORIGIN

Query Match 80.8%; Score 20.2; DB 9; Length 1052;
 Best Local Similarity 88.0%; Pred. No. 1.9e+02;
 Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAAGAGCGCAGACTGGTCTTCC 25
 |||||
 DB 715 AAAAAAGAGCGCAGCTGCTCTTCC 739

RESULT 45

AA138597 401 bp mRNA linear EST 11-FEB-1997
 LOCUS mg2c11.r1 StrataGene mouse heart (#937316) Mus musculus cDNA clone
 DEFINITION IMAGE:S86196 5', mRNA sequence.

ACCESSION AA138597
 VERSION AA138597.1 GI:1700809
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muroidea; Muridae; Murinae; Mus.

REFERENCE

1 (bases 1 to 401)
 Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
 Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
 Schellenberg, K., Seftoe, M., Tan, F., Underwood, K., Moore, B.,
 Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
 Waterston, R.

TITLE

The WashU-HMI Mouse EST Project

JOURNAL

COMMENT Unpublished (1996)
 Contact: Marra M/Mouse EST Project
 WashU-HMI Mouse EST Project
 Washington University School of MedicineP
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810

Email: mouseest@wustl.edu
 This clone is available royalty-free through LLNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.

MGI:360844

Seq primer: -28m13 rev1 ET from Amersham

High quality sequence stop: 389.

FEATURES

source

1..401
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="NIH Swiss"
 /db_xref="taxon:10090"
 /clone="IMAGE:586196"
 /sex="pooled"
 /tissue_type="heart"
 /dev_stage="13 day embryos"
 /lab_host="SOLR (kanamycin resistant)"
 /clone_lib="Stratagene mouse heart (#937316)"
 /note="Organ: heart; Vector: pBluescript SK-; Site 1:
 EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer:
 Oligo dt. 93 pooled NIH/Swiss 13 day embryo hearts.
 Average insert size: 1.0 kb; Uni-ZAP XR Vector; ~5'
 adaptor sequence: 5' GAATTCGACGACGAG 3' -3' adaptor
 sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'"

ORIGIN

Query Match 79.2%; Score 19.8; DB 1; Length 401;
 Best Local Similarity 91.3%; Pred. No. 2.6e+02;
 Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAAGAGCGCAGACTGGTCTT 23
 |||||
 DB 379 AAAAAAGAGCGCGACTGCTCTT 401

RESULT 46

LOCUS

DEFINITION

CO034709 868 bp mRNA linear EST 10-JUN-2004
 EST813093 Coccidioides posadasii spherule cDNA library, 0.5 to 5.3
 kb Coccidioides posadasii cDNA clone C1FBV39 5' end, mRNA sequence.

ACCESSION

CO034709

VERSION

CO034709.1 GI:48572066

KEYWORDS

SOURCE

ORGANISM

Coccidioides posadasii

REFERENCE

1 (bases 1 to 868)
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 Onygenales; mitosporic Onygenales; Coccidioides.

AUTHORS

Gardner, M.J. and Cole, G.T.

TITLE

Analysis of gene expression in Coccidioides posadasii mycelia and

JOURNAL

spherules via expressed sequence tags

COMMENT

Unpublished (2003)

Other ESTs: EST813092

Contact: Gardner MJ

The Institute for Genomic Research

9712 Medical Center Drive, Rockville, MD 20850, USA

Tel: 301 838 3519

Fax: 301 838 0208

Email: gardner@tigr.org

Seq primer: M13 Reverse.

FEATURES

source

1..868
 /organism="Coccidioides posadasii"
 /mol_type="mRNA"
 /strain="C735"
 /db_xref="taxon:199306"
 /clone="C1FBV39"
 /dev_stage="spherules"
 /lab_host="E. coli DH10B, T1 phage resistant"
 /clone_lib="Coccidioides posadasii spherule cDNA library,
 0.5 to 5.3 kb"
 /note="Vector: pExpress 1; Site_1: Not I; Site_2: Eco RV;
 Coccidioides posadasii spherule cDNA library, 0.5 to 5.3
 kb"

ORIGIN

Query Match 79.2%; Score 19.8; DB 7; Length 868;
 Best Local Similarity 91.3%; Pred. No. 2.9e+02;
 Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAAGAGCGCAGACTGGTCTT 23
 |||||
 DB 24 AAAAAAGAGCGCGACTGGTCTT 46

RESULT 47

LOCUS

DEFINITION

AA590658 159 bp mRNA linear EST 16-SEP-1997
 vm24g12.r1 Knowles Solter mouse blastocyst B1 Mus musculus CDNA
 clone IMAGE:991174 5', mRNA sequence.

ACCESSION

AA590658

VERSION

AA590658.1 GI:2403971

KEYWORDS

SOURCE

ORGANISM

Mus musculus (house mouse)

REFERENCE

AUTHORS

Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muroidea; Muridae; Murinae; Mus.
 1 (bases 1 to 159)
 Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
 Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
 Schellenberg, K., Seftoe, M., Tan, F., Underwood, K., Moore, B.,
 Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
 Waterston, R.

TITLE

JOURNAL

COMMENT

The WashU-HMI Mouse EST Project
 Unpublished (1996)
 Contact: Marra M/Mouse EST Project
 WashU-HMI Mouse EST Project

Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810

Email: mouseest@watson.wustl.edu

This clone is available royalty-free through LML; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.

MGI:563454

Putative full length read
vector to vector length is 160.

FEATURES

source
1. .159
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J x DBA/2J F1"
/db_xref="taxon:10090"
/clone="IMAGE:991174"
/tissue_type="blastocyst"
/dev_stage="embryo (pre-implantation)"
/lab_host="DH10B"
/clone_lib="Knowles Solter mouse blastocyst B1"
/note="Organ: embryo; Vector: pSPORT; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally from mRNA prepared
from 800 blastocysts. Primer: SalI(dtr):
5'-CGGTCCACGTCGACCGTGTCTTTT-3'. cDNAs were
cloned into the NotI/SalI sites of a pSPORT vector (Life
Technologies). Two different size selections: B1 (larger
inserts) and B3."

ORIGIN

Query Match 77.6%; Score 19.4; DB 1; Length 159;
Best Local Similarity 95.2%; Pred. No. 3.5e+02;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 AAGAGCGCAGACTGGTCTTCC 25
|||||
DB 14 AAGAGCGCAGACTGGTCTTCC 34

RESULT 48

AA823597 277 bp mRNA linear EST 17-FEB-1998
LOCUS
DEFINITION
IMAGE:1246064 5', mRNA sequence.
AA823597
VERSION
KEYWORDS
SOURCE
Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

AUTHORS
Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
Waterston, R.

TITLE

JOURNAL
COMMENT
The WashU-HMI Mouse EST Project
Unpublished (1996)
Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LML; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:659752
Seq primer: -28ml3 rev2 ET from Amersham
High quality sequence stop: 241.

FEATURES

Location/Qualifiers

source

1. .277
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:1246064"
/sex="male"
/tissue_type="mammary gland"
/dev_stage="4 weeks"
/lab_host="DH10B"
/clone_lib="Soares mammary gland NbMMG"
/note="Organ: mammary gland; Vector: pT73D-Pac
(Pharmacia) with a modified polylinker; Site 1: Not I;
Site 2: Eco RI; 1st strand cDNA was primed with a Not I -
oligo(dtr) primer [5',
TGTTACCAATCTGAGTGGAGCGCGCGAATGCTTTT-3']
T 3']; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pT73 vector.
RNA provided by Dr. Minoru Ko, Wayne State Univ. Library
constructed and normalized by Bento Soares and M.Fatima
Bonaldo."

ORIGIN

Query Match 77.6%; Score 19.4; DB 1; Length 277;
Best Local Similarity 95.2%; Pred. No. 3.8e+02;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 AAGAGCGCAGACTGGTCTTCC 25

|||||

DB 243 AAGAGCGCAGACTGGTCTTCC 263

RESULT 49

CR109101/c
LOCUS
DEFINITION
Reverse strand read from insert in 5'HPRT insertion targeting and
chromosome engineering clone MHPN151g12, genomic survey sequence.

ACCESSION

CR109101.1 GI:49856516

VERSION

GSS; genome survey sequence; MICEP.

KEYWORDS

SOURCE

ORGANISM

Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

AUTHORS
Adams, D.J., Biggs, P.J., Cox, A.V., Davies, R.M., van der Weyden, L.,
Jonkers, J., Smith, J., Plumb, R.W., Taylor, R.G., Nishijima, I., Yu, Y.,
Rogers, J. and Bradley, A.

TITLE

JOURNAL
Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. http://www.sanger.ac.uk/MICEP

FEATURES

source

1. .292
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="MHPN151g12"
/clone_lib="MHPN"

ORIGIN

Query Match 77.6%; Score 19.4; DB 11; Length 292;
Best Local Similarity 95.2%; Pred. No. 3.8e+02;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 AAGAGCGCAGACTGGTCTTCC 25

|||||

DB 185 AAGAGCGCAGACTGGTCTTCC 165

RESULT 50

BE197804/c

LOCUS BE197804 343 bp mRNA linear EST 26-JUN-2000
DEFINITION us77e02.x1 NCI_CGAP_Mam4 Mus musculus cDNA clone IMAGE:3257210 3',
mRNA sequence.
ACCESSION BE197804
VERSION BE197804.1 GI:8709907
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 343)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-i@mail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Priscilla Furth
Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
image.llnl.gov/image/html/iresources.shtml

FEATURES
source MGI:1068574
Seq primer: -40UP from Gibco
High quality sequence stop: 342.
Location/Qualifiers
1..343
/organism="Mus musculus"
/mol_type="mRNA"
/strain="NMRI"
/db_xref="taxon:10090"
/clone="IMAGE:3257210"
/tissue_type="tumor, gross tissue"
/dev_stage="5 months"
/lab_host="DH10B"
/clone_lib="NCI_CGAP Mam4"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Library constructed by Life Technologies. Investigators
providing samples: Lothar Hennighausen/Priscilla Furth,
NIH Reference for transgenic model: Li et al., Cell Growth
and Differentiation 7, 3-11 (1996)."

ORIGIN
Query Match 77.6%; Score 19.4; DB 2; Length 343;
Best Local Similarity 95.2%; Pred. No. 3.9e+02;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 5 AAGAGCGCAGACTGGTCTTCC 25
|||||
Db 167 AAGAGCGCAGACTGGTCTTCC 147

Search completed: February 3, 2006, 22:02:48
Job time : 2960.67 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 3, 2006, 11:35:31 ; Search time 80.6667 Seconds
(without alignments)
550.897 Million cell updates/sec

Title: US-10-719-900-35

Perfect score: 25

Sequence: 1 aaaaagagcgactgtcttcc 25

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database :

Issued Patents NA:*

- 1: /cgn2_6/ptodata/1/ina/1 COMB.seq:*
- 2: /cgn2_6/ptodata/1/ina/5 COMB.seq:*
- 3: /cgn2_6/ptodata/1/ina/6A COMB.seq:*
- 4: /cgn2_6/ptodata/1/ina/6B COMB.seq:*
- 5: /cgn2_6/ptodata/1/ina/H COMB.seq:*
- 6: /cgn2_6/ptodata/1/ina/PCTUS COMB.seq:*
- 7: /cgn2_6/ptodata/1/ina/PP COMB.seq:*
- 8: /cgn2_6/ptodata/1/ina/RE COMB.seq:*
- 9: /cgn2_6/ptodata/1/ina/backfileseq.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	17.8	71.2	11176	3	US-09-465-978A-51
2	17.6	70.4	601	3	US-09-949-016-72599
3	17.6	70.4	1285	3	US-09-999-833A-528
4	17.6	70.4	1285	3	US-10-020-445A-528
5	17.6	70.4	2602	3	US-09-999-833A-525
6	17.6	70.4	2602	3	US-10-020-445A-525
7	17.6	70.4	2652	3	US-09-667-373-3
8	17.6	70.4	2652	3	US-10-305-413-3
9	17.6	70.4	3291	2	US-08-574-763-1
10	17.6	70.4	3377	3	US-09-819-989-1
11	17.6	70.4	3377	3	US-10-273-992-1
12	17.6	70.4	3377	3	US-10-681-222-1
13	17.6	70.4	19650	3	US-09-819-989-3
14	17.6	70.4	19650	3	US-10-273-992-3
15	17.6	70.4	19650	3	US-10-681-222-3
16	17.6	70.4	155266	3	US-09-949-016-13870
17	17.2	68.8	601	3	US-09-949-016-170047
18	17.2	68.8	13743	3	US-09-949-016-16538
19	17.6	70.4	501	3	US-09-199-637A-244
20	17.6	70.4	534	3	US-09-199-637A-246
21	17.6	70.4	601	3	US-09-949-016-74479
22	17.6	70.4	601	3	US-09-949-016-190379
23	17.6	70.4	42235	3	US-09-199-637A-1
24	17.6	70.4	69813	3	US-09-949-016-12455

c 25	17	68.0	69813	3	US-09-949-016-13905	Sequence 13905, A
c 26	17	68.0	69813	3	US-09-949-016-13906	Sequence 13906, A
c 27	17	68.0	69813	3	US-09-949-016-12861	Sequence 12861, A
c 28	17	68.0	79824	3	US-09-949-016-13919	Sequence 13919, A
c 29	17	68.0	148567	3	US-09-801-876B-3	Sequence 3, Appli
c 30	17	68.0	148567	3	US-10-254-869-3	Sequence 3, Appli
c 31	17	68.0	148567	3	US-10-667-442-3	Sequence 3, Appli
c 32	17	68.0	255679	3	US-09-949-016-17189	Sequence 17189, A
c 33	16.8	67.2	429	3	US-09-252-991A-14962	Sequence 14962, A
c 34	16.8	67.2	828	3	US-09-252-991A-15086	Sequence 15086, A
c 35	16.8	67.2	1347	3	US-09-252-991A-14827	Sequence 14827, A
c 36	16.8	67.2	146095	3	US-09-949-016-12872	Sequence 12872, A
c 37	16.8	67.2	146104	3	US-09-949-016-12339	Sequence 12339, A
c 38	16.6	66.4	601	3	US-09-949-016-135413	Sequence 35413, A
c 39	16.6	66.4	601	3	US-09-949-016-77995	Sequence 77995, A
c 40	16.6	66.4	601	3	US-09-949-016-196973	Sequence 196973, A
c 41	16.6	66.4	601	3	US-09-949-016-198113	Sequence 198113, A
c 42	16.6	66.4	1061	3	US-09-270-767-10689	Sequence 10689, A
c 43	16.6	66.4	1123	2	US-08-458-023B-3	Sequence 3, Appli
c 44	16.6	66.4	6676	3	US-09-573-080A-211	Sequence 211, App
c 45	16.6	66.4	9277	3	US-09-949-016-14017	Sequence 14017, A
c 46	16.6	66.4	13011	2	US-08-791-849A-14	Sequence 14, Appli
c 47	16.6	66.4	21168	3	US-09-949-016-12643	Sequence 12643, A
c 48	16.6	66.4	26334	3	US-09-949-016-17354	Sequence 17354, A
c 49	16.6	66.4	28843	3	US-09-949-016-17325	Sequence 17325, A
c 50	16.6	66.4	34531	3	US-09-949-016-14604	Sequence 14604, A
c 51	16.6	66.4	325034	3	US-09-949-016-14957	Sequence 14957, A
c 52	16.6	66.4	389504	3	US-09-949-016-11774	Sequence 11774, A
c 53	16.6	66.4	1230025	3	US-09-198-452A-1	Sequence 1, Appli
c 54	16.6	66.4	1230230	3	US-09-438-185A-1	Sequence 36, Appli
c 55	16.2	64.8	491	2	US-08-133-711-36	Sequence 2135, Ap
c 56	16.2	64.8	601	3	US-09-949-002-21335	Sequence 2136, Ap
c 57	16.2	64.8	601	3	US-09-949-002-2136	Sequence 8498, Ap
c 58	16.2	64.8	601	3	US-09-949-002-8498	Sequence 8499, Ap
c 59	16.2	64.8	601	3	US-09-949-002-8499	Sequence 31, Appli
c 60	16.2	64.8	718	3	US-09-154-083-31	Sequence 3, Appli
c 61	16.2	64.8	1008	3	US-09-101-423B-3	Sequence 4, Appli
c 62	16.2	64.8	1187	2	US-08-474-177-4	Sequence 4, Appli
c 63	16.2	64.8	1187	2	US-08-487-033-4	Sequence 4, Appli
c 64	16.2	64.8	1187	2	US-08-480-810-4	Sequence 4, Appli
c 65	16.2	64.8	1187	2	US-08-508-735-4	Sequence 4, Appli
c 66	16.2	64.8	1187	2	US-08-848-251-4	Sequence 4, Appli
c 67	16.2	64.8	1187	2	US-08-486-047-4	Sequence 4, Appli
c 68	16.2	64.8	1187	2	US-09-120-130-4	Sequence 4, Appli
c 69	16.2	64.8	1187	3	US-09-115-252-4	Sequence 4, Appli
c 70	16.2	64.8	1187	3	US-08-986-515-4	Sequence 4, Appli
c 71	16.2	64.8	1187	3	US-09-120-128-4	Sequence 4, Appli
c 72	16.2	64.8	1187	3	US-09-120-129-4	Sequence 4, Appli
c 73	16.2	64.8	1187	3	US-09-201-139-4	Sequence 4, Appli
c 74	16.2	64.8	1187	3	US-09-120-131-4	Sequence 4, Appli
c 75	16.2	64.8	1259	3	US-09-809-545A-40	Sequence 40, Appli
c 76	16.2	64.8	1696	3	US-10-623-629-4	Sequence 4, Appli
c 77	16.2	64.8	3205	3	US-10-104-047-1412	Sequence 1412, Ap
c 78	16.2	64.8	3205	3	US-09-061-768A-3	Sequence 3, Appli
c 79	16.2	64.8	3205	3	US-09-764-246-3	Sequence 3, Appli
c 80	16.2	64.8	3605	3	US-09-098-901-1	Sequence 1, Appli
c 81	16.2	64.8	5039	3	US-09-386-816C-1	Sequence 1, Appli
c 82	16.2	64.8	5039	3	US-10-320-176-1	Sequence 1, Appli
c 83	16.2	64.8	5342	3	US-09-807-757C-1	Sequence 3, Appli
c 84	16.2	64.8	5889	3	US-09-402-929-3	Sequence 3, Appli
c 85	16.2	64.8	5889	3	US-09-660-299-1	Sequence 1, Appli
c 86	16.2	64.8	9196	3	US-09-780-175-96	Sequence 96, Appli
c 87	16.2	64.8	12537	2	US-09-971-773-3	Sequence 3, Appli
c 88	16.2	64.8	12537	3	US-08-611-280-4	Sequence 4, Appli
c 89	16.2	64.8	12537	3	US-09-195-940-4	Sequence 4, Appli
c 90	16.2	64.8	14507	3	US-09-562-466-4	Sequence 1, Appli
c 91	16.2	64.8	14507	3	US-08-785-150-1	Sequence 1, Appli
c 92	16.2	64.8	14507	3	US-09-435-377-1	Sequence 1, Appli
c 93	16.2	64.8	14507	3	US-09-973-928-1	Sequence 3, Appli
c 94	16.2	64.8	17056	3	US-09-345-041-3	Sequence 3, Appli
c 95	16.2	64.8	17056	3	US-09-358-055B-3	Sequence 3, Appli
c 96	16.2	64.8	17056	3	US-09-893-238-3	Sequence 3, Appli
c 97	16.2	64.8	24593	3	US-09-949-016-13433	Sequence 13433, A

```

C 98 16.2 64.8 26700 2 US-08-472-217-1 Sequence 1, Appli
C 99 16.2 64.8 26700 2 US-08-488-199-5 Sequence 5, Appli
C 100 16.2 64.8 26700 3 US-08-760-534A-1 Sequence 1, Appli
C 101 16.2 64.8 26700 3 US-09-336-757-1 Sequence 1, Appli
C 102 16.2 64.8 30310 3 US-09-657-346A-96 Sequence 96, Appli
C 103 16.2 64.8 90050 3 US-09-245-041-5 Sequence 5, Appli
C 104 16.2 64.8 90050 3 US-09-358-055B-5 Sequence 5, Appli
C 105 16.2 64.8 90050 3 US-09-893-238-5 Sequence 5, Appli
C 106 16.2 64.8 90923 3 US-09-949-002-623 Sequence 623, App
C 107 16.2 64.8 90925 3 US-09-949-002-789 Sequence 789, App
C 108 16.2 64.8 134008 3 US-09-949-016-13841 Sequence 13841, A
C 109 16.2 64.8 264358 3 US-09-949-016-15725 Sequence 15725, A
C 110 16 64.0 319 3 US-09-513-998C-35607 Sequence 35607, A
C 111 16 64.0 441 3 US-09-513-999C-648 Sequence 648, App
C 112 16 64.0 601 3 US-09-949-016-18039 Sequence 18039, A
C 113 16 64.0 601 3 US-09-949-016-19100 Sequence 19100, A
C 114 16 64.0 601 3 US-09-949-016-19101 Sequence 19101, A
C 115 16 64.0 601 3 US-09-949-016-59814 Sequence 59814, A
C 116 16 64.0 601 3 US-09-949-016-59815 Sequence 59815, A
C 117 16 64.0 601 3 US-09-949-016-59816 Sequence 59816, A
C 118 16 64.0 601 3 US-09-949-016-59817 Sequence 59817, A
C 119 16 64.0 601 3 US-09-949-016-68244 Sequence 68244, A
C 120 16 64.0 601 3 US-09-949-016-68245 Sequence 68245, A
C 121 16 64.0 601 3 US-09-949-016-68246 Sequence 68246, A
C 122 16 64.0 601 3 US-09-949-016-70266 Sequence 70266, A
C 123 16 64.0 601 3 US-09-949-016-145149 Sequence 145149,
C 124 16 64.0 601 3 US-09-949-016-157744 Sequence 157744,
C 125 16 64.0 601 3 US-09-949-016-157745 Sequence 157745,
C 126 16 64.0 625 3 US-09-533-559-3038 Sequence 3038, Ap
C 127 16 64.0 944 3 US-09-270-767-15167 Sequence 15167, A
C 128 16 64.0 1163 3 US-09-160-588-1 Sequence 1, Appli
C 129 16 64.0 1750 3 US-09-684-403-3 Sequence 3, Appli
C 130 16 64.0 1833 2 US-08-365-470-2 Sequence 2, Appli
C 131 16 64.0 2886 3 US-09-774-528-165 Sequence 165, App
C 132 16 64.0 2886 3 US-10-120-988-165 Sequence 165, App
C 133 16 64.0 3834 3 US-09-209-668-18 Sequence 18, Appli
C 134 16 64.0 3834 3 US-09-949-016-71 Sequence 71, Appli
C 135 16 64.0 3854 2 US-08-365-470-1 Sequence 1, Appli
C 136 16 64.0 3858 2 US-08-344-155C-98 Sequence 98, Appli
C 137 16 64.0 3858 3 US-09-009-490A-88 Sequence 88, Appli
C 138 16 64.0 3858 3 US-09-982-262C-89 Sequence 89, Appli
C 139 16 64.0 3862 3 US-09-949-016-4401 Sequence 4401, Ap
C 140 16 64.0 3863 3 US-08-482-073-1 Sequence 1, Appli
C 141 16 64.0 3863 9 5217870-1 Patent No. 5217870
C 142 16 64.0 4286 3 US-08-227-800A-2 Sequence 2, Appli
C 143 16 64.0 4286 3 US-08-921-954-2 Sequence 2, Appli
C 144 16 64.0 6102 3 US-09-949-016-2007 Sequence 2007, Ap
C 145 16 64.0 10095 3 US-09-949-016-16141 Sequence 16141, A
C 146 16 64.0 14566 3 US-09-949-016-16765 Sequence 16765, A
C 147 16 64.0 15384 3 US-09-949-016-11813 Sequence 11813, A
C 148 16 64.0 15413 3 US-09-949-016-16143 Sequence 16143, A
C 149 16 64.0 18149 3 US-09-949-016-13204 Sequence 13204, A
C 150 16 64.0 23849 3 US-09-949-016-13813 Sequence 13813, A
```

ALIGNMENTS

```

RESULT 1
US-09-465-978A-51
; Sequence 51, Application US/09465978A
; Patent No. 6867348
; GENERAL INFORMATION:
; APPLICANT: Zhang, Nine and Anthony Purchio
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR SCREENING FOR ANGIOGENESIS
; FILE REFERENCE: EXE-012.US
; CURRENT APPLICATION NUMBER: US/09/465, 978A
; CURRENT FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: 60/152,522
; PRIOR FILING DATE: 1999-09-03
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn Ver. 2.0
```

```

; SEQ ID NO 51
; LENGTH: 11176
; TYPE: DNA
; ORGANISM: Mus sp.
US-09-465-978A-51
Query Match 71.2%; Score 17.8; DB 3; Length 11176;
Best Local Similarity 90.5%; Pred. No. 1,1e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 5 AAGAGCGCAGACTGCTCTTCC 25
|||||
Db 10526 AAGACACAGACTGCTCTTCC 10546
|||||
RESULT 2
US-09-949-016-72599
; Sequence 72599, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 72599
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-72599
Query Match 70.4%; Score 17.6; DB 3; Length 601;
Best Local Similarity 83.3%; Pred. No. 86;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 AAAAAAGAGCGCAGACTGCTCTTC 24
|||||
Db 573 AAAAAAGAGCATAGACTGCTTCC 596
|||||
RESULT 3
US-09-999-833A-528/c
; Sequence 528, Application US/0999833A
; Patent No. 6916648
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
```

APPLICANT: Pan, James;
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Shelton, David L.
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
TITLE OF INVENTION: Acids Encoding the Same
FILE REFERENCE: P2630PIC65
CURRENT APPLICATION NUMBER: US/09/999,833A
CURRENT FILING DATE: 2001-10-24
PRIOR APPLICATION NUMBER: 09/918585
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064249
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/077450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077641
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077791
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/078004
PRIOR FILING DATE: 1998-03-13
PRIOR APPLICATION NUMBER: 60/078886
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078936
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078939
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079656
PRIOR FILING DATE: 1998-03-26
PRIOR APPLICATION NUMBER: 60/079664
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079689
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079663
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079786
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079920
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/079923
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/080105
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080107
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080165
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080194
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080327
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080328
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080333
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080334
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/081070
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081049
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081071
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081195
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081203
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081229
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081955
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081817
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081819
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081952
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081838
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082568
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082569
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082704
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082804
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082700
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082797
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082796
PRIOR FILING DATE: 1998-04-23
PRIOR APPLICATION NUMBER: 60/083336
PRIOR FILING DATE: 1998-04-27
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083392
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083495
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083496
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083499
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083545
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083554
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083558
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083559
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083500
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083742
PRIOR FILING DATE: 1998-04-30
PRIOR APPLICATION NUMBER: 60/084366
PRIOR FILING DATE: 1998-05-05
PRIOR APPLICATION NUMBER: 60/084414
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084441
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084637
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084639

```

; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084640
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084598
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084627
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084643
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085339
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085338
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085582
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085700
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085689
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085580
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085573
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697

```

Query Match 70.4%; Score 17.6; DB 3; Length 1285;

Best Local Similarity 83.3%; Pred. No. 97;

Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

QY 2 AAGAGGCGAGCTGCTTCC 25
   ||||| ||||| ||||| ||
Db 816 AGAAGAGCCGAGCTGCTCC 793

```

RESULT 4

```

US-10-020-445A-528/c
; Sequence 528, Application US/10020445A
; Patent No. 6962797
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James;
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey

```

```

; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2630PIC74
; CURRENT APPLICATION NUMBER: US/10/020.445A
; CURRENT FILING DATE: 2001-10-24
; PRIOR APPLICATION NUMBER: 09/918585
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/064249
; PRIOR FILING DATE: 1997-11-03
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066364
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/077450
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: 60/077632
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077641
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077649
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077791
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/078004
; PRIOR FILING DATE: 1998-03-13
; PRIOR APPLICATION NUMBER: 60/078886
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078936
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078939
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079664
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079663
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079786
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079920
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/079923
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/080105
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080107
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080165
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080194
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080327
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080328
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080333
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080334
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/081070
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081049

```

;
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081071
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081195
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081203
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081229
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081955
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081817
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081819
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081952
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081838
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/082568
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082569
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082704
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082804
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082700
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082797
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082796
; PRIOR FILING DATE: 1998-04-23
; PRIOR APPLICATION NUMBER: 60/083336
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/083392
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083495
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083496
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083499
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083545
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083554
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083558
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083559
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083500
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083742
; PRIOR FILING DATE: 1998-04-30
; PRIOR APPLICATION NUMBER: 60/084366
; PRIOR FILING DATE: 1998-05-05
; PRIOR APPLICATION NUMBER: 60/084414
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084441
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084637
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084639
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084640
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084598
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07

;
; PRIOR APPLICATION NUMBER: 60/084627
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084643
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085339
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085338
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085582
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085700
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085689
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085580
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085573
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697
;
Query Match 70.4%; Score 17.6; DB 3; Length 1285;
Best Local Similarity 83.3%; Pred. No. 97;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AAAAAAGCGCAGACTGGTCTTCC 25
| | | | | | | | | | | | | | | | | | | | | |
DB 816 AGAAGAGCCAGACTGGTCCACC 793

RESULT 5

US-09-999-833A-525/c
; Sequence 525, Application US/09999833A
; Patent No. 6916648
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Knapier, Mary A.
; APPLICANT: Pan, James;
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2630PIC65
; CURRENT APPLICATION NUMBER: US/09/999, 833A
; CURRENT FILING DATE: 2001-10-24
; PRIOR APPLICATION NUMBER: 09/918585

1	PRIOR FILING DATE: 2001-07-30	2	PRIOR APPLICATION NUMBER: 60/081222
3	PRIOR APPLICATION NUMBER: 60/062250	4	PRIOR FILING DATE: 1998-04-09
5	PRIOR FILING DATE: 1997-10-17	6	PRIOR APPLICATION NUMBER: 60/081955
7	PRIOR APPLICATION NUMBER: 60/064249	8	PRIOR FILING DATE: 1998-04-15
9	PRIOR FILING DATE: 1997-11-03	10	PRIOR APPLICATION NUMBER: 60/081817
11	PRIOR APPLICATION NUMBER: 60/065311	12	PRIOR FILING DATE: 1998-04-15
13	PRIOR FILING DATE: 1997-11-13	14	PRIOR APPLICATION NUMBER: 60/081819
15	PRIOR APPLICATION NUMBER: 60/066364	16	PRIOR FILING DATE: 1998-04-15
17	PRIOR FILING DATE: 1997-11-21	18	PRIOR APPLICATION NUMBER: 60/081952
19	PRIOR APPLICATION NUMBER: 60/077450	20	PRIOR FILING DATE: 1998-04-15
21	PRIOR FILING DATE: 1998-03-10	22	PRIOR APPLICATION NUMBER: 60/081838
23	PRIOR APPLICATION NUMBER: 60/077632	24	PRIOR FILING DATE: 1998-04-15
25	PRIOR FILING DATE: 1998-03-11	26	PRIOR APPLICATION NUMBER: 60/082568
27	PRIOR APPLICATION NUMBER: 60/077641	28	PRIOR FILING DATE: 1998-04-21
29	PRIOR FILING DATE: 1998-03-11	30	PRIOR APPLICATION NUMBER: 60/082569
31	PRIOR APPLICATION NUMBER: 60/077649	32	PRIOR FILING DATE: 1998-04-21
33	PRIOR FILING DATE: 1998-03-11	34	PRIOR APPLICATION NUMBER: 60/082704
35	PRIOR APPLICATION NUMBER: 60/077791	36	PRIOR FILING DATE: 1998-04-22
37	PRIOR FILING DATE: 1998-03-12	38	PRIOR APPLICATION NUMBER: 60/082804
39	PRIOR APPLICATION NUMBER: 60/078004	40	PRIOR FILING DATE: 1998-04-22
41	PRIOR FILING DATE: 1998-03-13	42	PRIOR APPLICATION NUMBER: 60/082700
43	PRIOR APPLICATION NUMBER: 60/078886	44	PRIOR FILING DATE: 1998-04-22
45	PRIOR FILING DATE: 1998-03-20	46	PRIOR APPLICATION NUMBER: 60/082797
47	PRIOR APPLICATION NUMBER: 60/078936	48	PRIOR FILING DATE: 1998-04-22
49	PRIOR FILING DATE: 1998-03-20	50	PRIOR APPLICATION NUMBER: 60/082796
51	PRIOR APPLICATION NUMBER: 60/078910	52	PRIOR FILING DATE: 1998-04-23
53	PRIOR FILING DATE: 1998-03-20	54	PRIOR APPLICATION NUMBER: 60/083336
55	PRIOR APPLICATION NUMBER: 60/078939	56	PRIOR FILING DATE: 1998-04-27
57	PRIOR FILING DATE: 1998-03-20	58	PRIOR APPLICATION NUMBER: 60/083322
59	PRIOR APPLICATION NUMBER: 60/079294	60	PRIOR FILING DATE: 1998-04-28
61	PRIOR FILING DATE: 1998-03-25	62	PRIOR APPLICATION NUMBER: 60/083392
63	PRIOR APPLICATION NUMBER: 60/079656	64	PRIOR FILING DATE: 1998-04-29
65	PRIOR FILING DATE: 1998-03-26	66	PRIOR APPLICATION NUMBER: 60/083495
67	PRIOR APPLICATION NUMBER: 60/079664	68	PRIOR FILING DATE: 1998-04-29
69	PRIOR FILING DATE: 1998-03-27	70	PRIOR APPLICATION NUMBER: 60/083496
71	PRIOR APPLICATION NUMBER: 60/079689	72	PRIOR FILING DATE: 1998-04-29
73	PRIOR FILING DATE: 1998-03-27	74	PRIOR APPLICATION NUMBER: 60/083499
75	PRIOR APPLICATION NUMBER: 60/079663	76	PRIOR FILING DATE: 1998-04-29
77	PRIOR FILING DATE: 1998-03-27	78	PRIOR APPLICATION NUMBER: 60/083545
79	PRIOR APPLICATION NUMBER: 60/079728	80	PRIOR FILING DATE: 1998-04-29
81	PRIOR FILING DATE: 1998-03-27	82	PRIOR APPLICATION NUMBER: 60/083554
83	PRIOR APPLICATION NUMBER: 60/079786	84	PRIOR FILING DATE: 1998-04-29
85	PRIOR FILING DATE: 1998-03-27	86	PRIOR APPLICATION NUMBER: 60/083558
87	PRIOR APPLICATION NUMBER: 60/079920	88	PRIOR FILING DATE: 1998-04-29
89	PRIOR FILING DATE: 1998-03-30	90	PRIOR APPLICATION NUMBER: 60/083559
91	PRIOR APPLICATION NUMBER: 60/079923	92	PRIOR FILING DATE: 1998-04-29
93	PRIOR FILING DATE: 1998-03-30	94	PRIOR APPLICATION NUMBER: 60/083500
95	PRIOR APPLICATION NUMBER: 60/080105	96	PRIOR FILING DATE: 1998-04-29
97	PRIOR FILING DATE: 1998-03-31	98	PRIOR APPLICATION NUMBER: 60/083742
99	PRIOR APPLICATION NUMBER: 60/080107	100	PRIOR FILING DATE: 1998-04-30
101	PRIOR FILING DATE: 1998-03-31	102	PRIOR APPLICATION NUMBER: 60/084366
103	PRIOR APPLICATION NUMBER: 60/080165	104	PRIOR FILING DATE: 1998-05-05
105	PRIOR FILING DATE: 1998-03-31	106	PRIOR APPLICATION NUMBER: 60/084414
107	PRIOR APPLICATION NUMBER: 60/080194	108	PRIOR FILING DATE: 1998-05-06
109	PRIOR FILING DATE: 1998-03-31	110	PRIOR APPLICATION NUMBER: 60/084441
111	PRIOR APPLICATION NUMBER: 60/080327	112	PRIOR FILING DATE: 1998-05-06
113	PRIOR FILING DATE: 1998-04-01	114	PRIOR APPLICATION NUMBER: 60/084637
115	PRIOR APPLICATION NUMBER: 60/080328	116	PRIOR FILING DATE: 1998-05-07
117	PRIOR FILING DATE: 1998-04-01	118	PRIOR APPLICATION NUMBER: 60/084639
119	PRIOR APPLICATION NUMBER: 60/080333	120	PRIOR FILING DATE: 1998-05-07
121	PRIOR FILING DATE: 1998-04-01	122	PRIOR APPLICATION NUMBER: 60/084640
123	PRIOR APPLICATION NUMBER: 60/080334	124	PRIOR FILING DATE: 1998-05-07
125	PRIOR FILING DATE: 1998-04-01	126	PRIOR APPLICATION NUMBER: 60/084598
127	PRIOR APPLICATION NUMBER: 60/081070	128	PRIOR FILING DATE: 1998-05-07
129	PRIOR FILING DATE: 1998-04-08	130	PRIOR APPLICATION NUMBER: 60/084600
131	PRIOR APPLICATION NUMBER: 60/081049	132	PRIOR FILING DATE: 1998-05-07
133	PRIOR FILING DATE: 1998-04-08	134	PRIOR APPLICATION NUMBER: 60/084627
135	PRIOR APPLICATION NUMBER: 60/081071	136	PRIOR FILING DATE: 1

```

; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085582
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085700
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085689
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085580
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085573
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697

Query Match 70.4%; Score 17.6; DB 3; Length 2602;
Best Local Similarity 83.3%; Pred. No. 1.1e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AAAAGAGCGCAGACTGGTCTTCC 25
Db 785 AGAAGAGCCAGACTGGTCCACC 762

RESULT 6
US-10-020-445A-525/c
Sequence 525, Application US/10020445A
Patent No. 6962797
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi
APPLICANT: Baker Kevin P.
APPLICANT: Botstein, David
APPLICANT: Deenoyers, Luc
APPLICANT: Baton, Dan
APPLICANT: Ferrara, Napoleon
APPLICANT: Filvaroff, Ellen
APPLICANT: Pong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth J.
APPLICANT: Kljavin, Ivar J.
APPLICANT: Kuo, Sophia S.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James;
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Shelton, David L.
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2630P1C74
CURRENT APPLICATION NUMBER: US/10/020,445A
CURRENT FILING DATE: 2001-10-24
PRIOR APPLICATION NUMBER: 09/918585
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064249
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066364
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/077450
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: 60/077632
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077641
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077649
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077791
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/078004
; PRIOR FILING DATE: 1998-03-13
; PRIOR APPLICATION NUMBER: 60/078886
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078936
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078939
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079664
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079663
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079786
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079920
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/079923
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/080105
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080107
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080165
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080194
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080327
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080328
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080333
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080334
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/081070
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081049
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081071
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081195
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081203
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081229
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081955
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081817
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081819
```

; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081952
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081838
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/082568
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082569
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082704
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082804
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082700
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/083336
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/083392
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083495
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083496
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083499
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083545
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083554
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083558
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083559
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083500
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083742
; PRIOR FILING DATE: 1998-04-30
; PRIOR APPLICATION NUMBER: 60/084366
; PRIOR FILING DATE: 1998-05-05
; PRIOR APPLICATION NUMBER: 60/084414
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084441
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084637
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084639
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084640
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084598
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084627
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084643
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085339
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085338
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085592
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085700
; PRIOR FILING DATE: 1998-05-15

; PRIOR APPLICATION NUMBER: 60/085689
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085580
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085573
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697
; PRIOR APPLICATION NUMBER: 60/085697
Query Match 70.4%; Score 17.6; DB 3; Length 2602;
Best Local Similarity 83.3%; Pred. No. 1.1e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 2 AAAAAGAGCGCAGACTGGTCTTCC 25
Db 785 AGAAAGAGCCAGACTGGTCCACC 762
RESULT 7
US-09-667-373-3/c
; Sequence 3, Application US/09667373
; Patent No. 6524840
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Nehls, Michael C.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 6524840el Human Endothelin Converting
; FILE REFERENCE: LEX-0043-USA
; CURRENT APPLICATION NUMBER: US/09/667,373
; CURRENT FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: US 60/156,102
; PRIOR FILING DATE: 1999-09-24
; PRIOR APPLICATION NUMBER: US 60/176,689
; PRIOR FILING DATE: 2000-01-18
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 2652
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-667-373-3
Query Match 70.4%; Score 17.6; DB 3; Length 2652;
Best Local Similarity 83.3%; Pred. No. 1.1e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 2 AAAAAGAGCGCAGACTGGTCTTCC 25
Db 1127 AGAAAGAGCCAGACTGGTCCACC 1104
RESULT 8
US-10-305-413-3/c
; Sequence 3, Application US/10305413
; Patent No. 6835564
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Nehls, Michael C.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 6835564el Human Endothelin Converting
; FILE REFERENCE: LEX-0043-USA
; CURRENT APPLICATION NUMBER: US/10/305,413


```
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3377
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-681-222-1

Query Match          70.4%; Score 17.6; DB 3; Length 3377;
Best Local Similarity 83.3%; Pred. No. 1.1e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY  2  AAAAAGAGCGCAGACTGGTCTTCC 25
Db    1024 AGAAAGAGCCAGACTGGTCCACC 1001

RESULT 13
US-09-819-989-3/c
; Sequence 3, Application US/09819989
; Patent No. 6482629
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al.
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001200
; CURRENT APPLICATION NUMBER: US/09/819,989
; CURRENT FILING DATE: 2001-03-29
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 19650
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(19650)
; OTHER INFORMATION: n = A,T,C or G
US-09-819-989-3

Query Match          70.4%; Score 17.6; DB 3; Length 19650;
Best Local Similarity 83.3%; Pred. No. 1.5e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY  2  AAAAAGAGCGCAGACTGGTCTTCC 25
Db    4443 AGAAAGAGCCAGACTGGTCCACC 4420

RESULT 14
US-10-273-992-3/c
; Sequence 3, Application US/10273992
; Patent No. 6664093
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al.
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001200DIV
; CURRENT APPLICATION NUMBER: US/10/273,992
; CURRENT FILING DATE: 2002-10-21
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 19650
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(19650)
; OTHER INFORMATION: n = A,T,C or G
US-10-273-992-3
```

```
Query Match          70.4%; Score 17.6; DB 3; Length 19650;
Best Local Similarity 83.3%; Pred. No. 1.5e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY  2  AAAAAGAGCGCAGACTGGTCTTCC 25
Db    4443 AGAAAGAGCCAGACTGGTCCACC 4420

RESULT 15
US-10-681-222-3/c
; Sequence 3, Application US/10681222
; Patent No. 6825025
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al.
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001200-DIV II
; CURRENT APPLICATION NUMBER: US/10/681,222
; CURRENT FILING DATE: 2003-10-09
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 19650
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(19650)
; OTHER INFORMATION: n = A,T,C or G
US-10-681-222-3
```

```
Query Match          70.4%; Score 17.6; DB 3; Length 19650;
Best Local Similarity 83.3%; Pred. No. 1.5e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY  2  AAAAAGAGCGCAGACTGGTCTTCC 25
Db    4443 AGAAAGAGCCAGACTGGTCCACC 4420

RESULT 16
US-09-949-016-13870/c
; Sequence 13870, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13870
; LENGTH: 155266
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13870

Query Match          70.4%; Score 17.6; DB 3; Length 155266;
Best Local Similarity 83.3%; Pred. No. 2.1e+02;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY  1  AAAAAGAGCGCAGACTGGTCTTC 24
```

Db 93329 AAAAAAGCATAGACTGGTTCC 93306
|||||||

RESULT 17
US-09-949-016-170047
; Sequence 170047, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 170047
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-170047

Query Match 68.8%; Score 17.2; DB 3; Length 601;
Best Local Similarity 86.4%; Pred. No. 1.3e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 AAAAAAGCGCAGACTGGTCT 22
|||||||
Db 322 AAAAAAGAAAGCAGTCTGGTCT 343

RESULT 18
US-09-949-016-16538/c
; Sequence 16538, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16538
; LENGTH: 13743
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-16538

Query Match 68.8%; Score 17.2; DB 3; Length 13743;
Best Local Similarity 86.4%; Pred. No. 2.2e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 AAAAAAGCGCAGACTGGTCT 22
|||||||
Db 4472 AAAAAAGAAAGCAGTCTGGTCT 4451

RESULT 19

US-09-199-637A-244
; Sequence 244, Application US/09199637A
; Patent No. 6355411
; GENERAL INFORMATION:
; APPLICANT: Ausubel, Frederick
; APPLICANT: Goodman, Howard M.
; APPLICANT: Rahme, Laurence G.
; APPLICANT: Mahajan-Miklos, Shalina
; APPLICANT: Tan, Man-Wah
; APPLICANT: Cao, Hui
; APPLICANT: Drenkard, Eliana
; APPLICANT: Tsongalis, John
; TITLE OF INVENTION: VIRULENCE-ASSOCIATED NUCLEIC ACID
; TITLE OF INVENTION: SEQUENCES AND USES THEREOF
; FILE REFERENCE: 00786/361002
; CURRENT APPLICATION NUMBER: US/09/199,637A
; CURRENT FILING DATE: 1998-11-25
; PRIOR APPLICATION NUMBER: 60/066,517
; PRIOR FILING DATE: 1997-11-25
; NUMBER OF SEQ ID NOS: 437
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 244
; LENGTH: 501
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-199-637A-244

Query Match 68.0%; Score 17; DB 3; Length 501;
Best Local Similarity 80.0%; Pred. No. 1.6e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 1 AAAAAAGCGCAGACTGGTCTTCC 25
|||||||
Db 253 AAAAAAGCGCGCCAGATCTTCC 277

RESULT 20
US-09-199-637A-246/c
; Sequence 246, Application US/09199637A
; Patent No. 6355411
; GENERAL INFORMATION:
; APPLICANT: Ausubel, Frederick
; APPLICANT: Goodman, Howard M.
; APPLICANT: Rahme, Laurence G.
; APPLICANT: Mahajan-Miklos, Shalina
; APPLICANT: Tan, Man-Wah
; APPLICANT: Cao, Hui
; APPLICANT: Drenkard, Eliana
; APPLICANT: Tsongalis, John
; TITLE OF INVENTION: VIRULENCE-ASSOCIATED NUCLEIC ACID
; TITLE OF INVENTION: SEQUENCES AND USES THEREOF
; FILE REFERENCE: 00786/361002
; CURRENT APPLICATION NUMBER: US/09/199,637A
; CURRENT FILING DATE: 1998-11-25
; PRIOR APPLICATION NUMBER: 60/066,517
; PRIOR FILING DATE: 1997-11-25
; NUMBER OF SEQ ID NOS: 437
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 246
; LENGTH: 534
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-199-637A-246

Query Match 68.0%; Score 17; DB 3; Length 534;
Best Local Similarity 80.0%; Pred. No. 1.6e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 1 AAAAAAGCGCAGACTGGTCTTCC 25
|||||||
Db 297 AAAAAAGCGCGCCAGATCTTCC 273

```
RESULT 21
US-09-949-016-74479/c
; Sequence 74479, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 74479
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-74479

Query Match      68.0%; Score 17; DB 3; Length 601;
Best Local Similarity 80.0%; Pred. No. 1.7e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAAAGAGCGCAGACTGGTCTTCC 25
Db 231 AAAAAAGAGAGAGACTGGTCTTAC 207

RESULT 22
US-09-949-016-190379/c
; Sequence 190379, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 190379
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-190379

Query Match      68.0%; Score 17; DB 3; Length 601;
Best Local Similarity 80.0%; Pred. No. 1.7e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAAAGAGCGCAGACTGGTCTTCC 25
Db 139 AAAAAAAGGCAGACTGATCTGCC 115

RESULT 23
US-09-199-637A-1
; Sequence 1, Application US/09199637A
; Patent No. 6355411
; GENERAL INFORMATION:
; APPLICANT: Ausubel, Frederick
; APPLICANT: Goodman, Howard M.
; APPLICANT: Rahme, Laurence G.
; APPLICANT: Mahajan-Miklos, Shalina
; APPLICANT: Tan, Man-Wah
; APPLICANT: Cao, Hui
; APPLICANT: Drenkard, Eliana
; APPLICANT: Tsongalis, John
; TITLE OF INVENTION: VIRULENCE-ASSOCIATED NUCLEIC ACID
; SEQUENCES AND USES THEREOF
; FILE REFERENCE: 00786/361002
; CURRENT APPLICATION NUMBER: US/09/199,637A
; CURRENT FILING DATE: 1998-11-25
; PRIOR APPLICATION NUMBER: 60/066,517
; PRIOR FILING DATE: 1997-11-25
; NUMBER OF SEQ ID NOS: 437
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 42235
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; FEATURE:
; NAME/KEY: variation
; LOCATION: (1)...(42235)
; OTHER INFORMATION: N is any nucleic acid.
US-09-199-637A-1

Query Match      68.0%; Score 17; DB 3; Length 42235;
Best Local Similarity 80.0%; Pred. No. 3.3e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAAAGAGCGCAGACTGGTCTTCC 25
Db 18731 AAAAAAAGCGCCGACAGATCTTCC 18755

RESULT 24
US-09-949-016-12455/c
; Sequence 12455, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12455
; LENGTH: 69813
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12455

Query Match      68.0%; Score 17; DB 3; Length 69813;
Best Local Similarity 80.0%; Pred. No. 3.5e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAAAGAGCGCAGACTGGTCTTCC 25
Db 4997 AAAAAAAGCTCAAACTTGCCTTCC 4973

RESULT 25
US-09-949-016-13905/c
; Sequence 13905, Application US/09949016
```

```
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13905
; LENGTH: 69813
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13905

Query Match          68.0%; Score 17; DB 3; Length 69813;
Best Local Similarity 80.0%; Pred. No. 3.5e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAAAGCGCAGACTGGTCTTCC 25
    ||||| ||| ||| ||| ||| ||| |||
DB 4997 AAAAAAAGCTCAAACTTGCCTTCC 4973

RESULT 26
US-09-949-016-13906/c
; Sequence 13906, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13906
; LENGTH: 69813
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13906

Query Match          68.0%; Score 17; DB 3; Length 69813;
Best Local Similarity 80.0%; Pred. No. 3.5e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAAAGCGCAGACTGGTCTTCC 25
    ||||| ||| ||| ||| ||| ||| |||
DB 4997 AAAAAAAGCTCAAACTTGCCTTCC 4973

RESULT 27
US-09-949-016-12861/c
; Sequence 12861, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12861
; LENGTH: 69813
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12861

Query Match          68.0%; Score 17; DB 3; Length 69813;
Best Local Similarity 80.0%; Pred. No. 3.5e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAAAGCGCAGACTGGTCTTCC 25
    ||||| ||| ||| ||| ||| ||| |||
DB 2173 AAAAAAGAGAAAGACTGGTCTTAC 2149

RESULT 29
US-09-801-876B-3/c
; Sequence 3, Application US/09801876B
; Patent No. 6492155
; GENERAL INFORMATION:
; APPLICANT: YE, Jane et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; FILE REFERENCE: CL001160
; CURRENT APPLICATION NUMBER: US/09/801,876B
; CURRENT FILING DATE: 2001-03-09
; NUMBER OF SEQ ID NOS: 8
```

```
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12861
; LENGTH: 69833
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12861

Query Match          68.0%; Score 17; DB 3; Length 69833;
Best Local Similarity 80.0%; Pred. No. 3.5e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAAAGCGCAGACTGGTCTTCC 25
    ||||| ||| ||| ||| ||| ||| |||
DB 5017 AAAAAAAGCTCAAACTTGCCTTCC 4993

RESULT 28
US-09-949-016-13919/c
; Sequence 13919, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13919
; LENGTH: 79824
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13919

Query Match          68.0%; Score 17; DB 3; Length 79824;
Best Local Similarity 80.0%; Pred. No. 3.5e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAAAGCGCAGACTGGTCTTCC 25
    ||||| ||| ||| ||| ||| ||| |||
DB 2173 AAAAAAGAGAAAGACTGGTCTTAC 2149

RESULT 29
US-09-801-876B-3/c
; Sequence 3, Application US/09801876B
; Patent No. 6492155
; GENERAL INFORMATION:
; APPLICANT: YE, Jane et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; FILE REFERENCE: CL001160
; CURRENT APPLICATION NUMBER: US/09/801,876B
; CURRENT FILING DATE: 2001-03-09
; NUMBER OF SEQ ID NOS: 8
```

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 148567
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(148567)
; OTHER INFORMATION: n = A,T,C or G
US-09-801-876B-3
```

```
Query Match          68.0%; Score 17; DB 3; Length 148567;
Best Local Similarity 80.0%; Pred. No. 3.9e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
```

```
QY 1 AAAAAAGCGCAGACTGCTCTTCC 25
    ||||| ||||| ||||| |||||
Db 44191 AAAAAAGCTAAAGCTGTTCTTCC 44167
```

RESULT 30

```
US-10-254-869-3/c
; Sequence 3, Application US/10254869
; Patent No. 6653117
; GENERAL INFORMATION:
; APPLICANT: YE, Jane et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001160DIV
; CURRENT APPLICATION NUMBER: US/10/254,869
; CURRENT FILING DATE: 2002-09-26
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 148567
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(148567)
; OTHER INFORMATION: n = A,T,C or G
US-10-254-869-3
```

```
Query Match          68.0%; Score 17; DB 3; Length 148567;
Best Local Similarity 80.0%; Pred. No. 3.9e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
```

```
QY 1 AAAAAAGCGCAGACTGCTCTTCC 25
    ||||| ||||| ||||| |||||
Db 44191 AAAAAAGCTAAAGCTGTTCTTCC 44167
```

RESULT 31

```
US-10-667-442-3/c
; Sequence 3, Application US/10667442
; Patent No. 6821765
; GENERAL INFORMATION:
; APPLICANT: YE, Jane et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001160DIV II
; CURRENT APPLICATION NUMBER: US/10/667,442
; CURRENT FILING DATE: 2003-09-23
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 148567
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
```

```
; LOCATION: (1)...(148567)
; OTHER INFORMATION: n = A,T,C or G
US-10-667-442-3
```

```
Query Match          68.0%; Score 17; DB 3; Length 148567;
Best Local Similarity 80.0%; Pred. No. 3.9e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
```

```
QY 1 AAAAAAGCGCAGACTGCTCTTCC 25
    ||||| ||||| ||||| |||||
Db 44191 AAAAAAGCTAAAGCTGTTCTTCC 44167
```

RESULT 32

```
US-09-949-016-17189/c
; Sequence 17189, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17189
; LENGTH: 255679
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(255679)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17189
```

```
Query Match          68.0%; Score 17; DB 3; Length 255679;
Best Local Similarity 80.0%; Pred. No. 4.2e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
```

```
QY 1 AAAAAAGCGCAGACTGCTCTTCC 25
    ||||| ||||| ||||| |||||
Db 236732 AAAAAAGCGCAGACTGCTGCC 236708
```

RESULT 33

```
US-09-252-991A-14962
; Sequence 14962, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 14962
; LENGTH: 429
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-14962
```

```
Query Match          67.2%; Score 16.8; DB 3; Length 429;
Best Local Similarity 90.0%; Pred. No. 2e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 AAAGAGCGCAGACTGGTCTT 23
DB 273 AAGGAGCGCAGACTGGTCTAT 292

RESULT 34
US-09-252-991A-15086
; Sequence 15086, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 15086
; LENGTH: 828
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-15086

Query Match          67.2%; Score 16.8; DB 3; Length 828;
Best Local Similarity 90.0%; Pred. No. 2.2e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 AAAGAGCGCAGACTGGTCTT 23
DB 344 AAGGAGCGCAGACTGGTCTAT 363

RESULT 35
US-09-252-991A-14827
; Sequence 14827, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 14827
; LENGTH: 1347
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-14827

Query Match          67.2%; Score 16.8; DB 3; Length 1347;
Best Local Similarity 90.0%; Pred. No. 2.3e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 AAAGAGCGCAGACTGGTCTT 23
DB 562 AAGGAGCGCAGACTGGTCTAT 581

RESULT 36
US-09-949-016-12872/c
; Sequence 12872, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12872
; LENGTH: 146095
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12872

Query Match          67.2%; Score 16.8; DB 3; Length 146095;
Best Local Similarity 90.0%; Pred. No. 4.8e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAAGAGCGCAGACTGGT 20
DB 66917 AGAAACAGCGCAGACTGGT 66898

RESULT 37
US-09-949-016-13239/c
; Sequence 13239, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13239
; LENGTH: 146104
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13239

Query Match          67.2%; Score 16.8; DB 3; Length 146104;
Best Local Similarity 90.0%; Pred. No. 4.8e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAAGAGCGCAGACTGGT 20
DB 66917 AGAAACAGCGCAGACTGGT 66898

RESULT 38
US-09-949-016-35413/c
; Sequence 35413, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
```

```
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 35413
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-35413

Query Match          66.4%; Score 16.6; DB 3; Length 601;
Best Local Similarity 82.6%; Pred. No. 2.6e+02;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAAGCGCAGACTGGTCTT 23
   ||||| ||||| ||||| ||||| |||||
Db 236 AAAAAAAGCGCAGACTTGT 214

RESULT 39
US-09-949-016-77995/c
; Sequence 77995, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 77995
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-77995

Query Match          66.4%; Score 16.6; DB 3; Length 601;
Best Local Similarity 82.6%; Pred. No. 2.6e+02;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAAGCGCAGACTGGTCTT 23
   ||||| ||||| ||||| ||||| |||||
Db 236 AAAAAAAGCGCAGACTTGT 214

RESULT 40
US-09-949-016-196973/c
; Sequence 196973, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
```

```
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 196973
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-196973

Query Match          66.4%; Score 16.6; DB 3; Length 601;
Best Local Similarity 82.6%; Pred. No. 2.6e+02;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 AAAAGAGCGCAGACTGGTCTTC 25
   ||| ||||| ||||| ||||| |||||
Db 499 AAACGCTGCAGACTGGTCTTC 477

RESULT 41
US-09-949-016-198113/c
; Sequence 198113, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 198113
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-198113

Query Match          66.4%; Score 16.6; DB 3; Length 601;
Best Local Similarity 82.6%; Pred. No. 2.6e+02;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 AAAAGAGCGCAGACTGGTCTTC 25
   ||| ||||| ||||| ||||| |||||
Db 188 AAAATAGCCAGACTGTTTTC 166

RESULT 42
US-09-270-767-10689/c
; Sequence 10689, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10689
; LENGTH: 1061
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-10689
```


QY 1 AAAAAAGCGCAGACTGGTCTT 23
||||| ||||| ||||| |||||
Db 3212 AAAAAAGCGCAGACTGGTTT 3190

RESULT 46
US-08-791-849A-14/c
; Sequence 14, Application US/08791849A
; Patent No. 5914449
; GENERAL INFORMATION:
; APPLICANT: Makoto MURASE et al.
; TITLE OF INVENTION: Method for Increasing Storage
; TITLE OF INVENTION: Lipid Content in Plant Seed
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/791,849A
; FILING DATE: January 30, 1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek, Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX:
; TELEX:

INFORMATION FOR SEQ ID NO: 14:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 13011 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: rat (Rattus norvegicus)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: Join(3212..3218, 3766..3948, 5917..6008,
; LOCATION: 6152..6283, 6418..6604, 6921..7191, 7302..7452, 7665..7817, 7911..8077
; LOCATION: 9298..9479, 10163..10269)
; FEATURE:
; NAME/KEY: exon
; LOCATION: Join(3194..3218, 3766..3948, 5917..6008,
; LOCATION: 6152..6283, 6418..6604, 6921..7191, 7302..7452, 7665..7817, 7911..8077
; FEATURE:
; NAME/KEY: intron
; LOCATION: Join(3219..3765, 3949..5916, 6009..6151,
; LOCATION: 6284..6417, 6605..6920, 7191..7301, 7453..7664, 7818..7910, 8078..9297
; LOCATION: 9480..10162)
US-08-791-849A-14

Query Match 66.4%; Score 16.6; DB 2; Length 13011;
Best Local Similarity 82.8%; Pred. No. 4.2e+02;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 AAAAAAGCGCAGACTGGTCTTCC 25
||||| ||||| ||||| |||||

Db 9788 ATAAGAGCACTGACTGTTCTTCC 9766

RESULT 47

US-09-949-016-12643/c
; Sequence 12643, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12643
; LENGTH: 21168
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12643

Query Match 66.4%; Score 16.6; DB 3; Length 21168;
Best Local Similarity 82.6%; Pred. No. 4.5e+02;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAAGCGCAGACTGGTCTT 23
||||| ||||| ||||| |||||

Db 3214 AAAAAAGCGCAGACTGGTTT 3192

RESULT 48

US-09-949-016-17354/c
; Sequence 17354, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17354
; LENGTH: 26334
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17354

Query Match 66.4%; Score 16.6; DB 3; Length 26334;
Best Local Similarity 82.6%; Pred. No. 4.7e+02;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 AAAAAAGCGCAGACTGGTCTTCC 25
||||| ||||| ||||| |||||

Db 10129 AAAATAGCCAGACTGTTTCC 10107

RESULT 49

US-09-949-016-17325

Search completed: February 3, 2006, 16:32:55
Job time : 86.6667 secs

```
; Sequence 17325, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17325
; LENGTH: 28843
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(28843)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17325
```

```
Query Match 66.4%; Score 16.6; DB 3; Length 28843;
Best Local Similarity 82.6%; Pred. No. 4.7e+02;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
```

```
QY 3 AAAAGCGGAGACTGGTCTTCC 25
DB 9644 AAACGCTGGCAGACTGGTCTTCC 9666
```

```
RESULT 50
US-09-949-016-14604/c
; Sequence 14604, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14604
; LENGTH: 34531
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(34531)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14604
```

```
Query Match 66.4%; Score 16.6; DB 3; Length 34531;
Best Local Similarity 82.6%; Pred. No. 4.9e+02;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
```

```
QY 1 AAAAAGAGCGCAGACTGGTCTT 23
DB 17328 AAAAAGAGAGAGACTGGTCTT 17306
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 3, 2006, 14:37:28 ; Search time 365.556 Seconds
(without alignments)
565.535 Million cell updates/sec

Title: US-10-719-900-35

Perfect score: 25

Sequence: 1 aaaaaagcgcagactgtcttc 25

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database : Published Applications NA_Main:

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	25	100.0	25	8	US-10-719-900-35
2	23.4	93.6	25	8	US-10-719-900-36
3	22.4	89.6	25	8	US-10-719-900-342
4	21.8	87.2	36599	7	US-10-322-281-679
5	20.8	83.2	25	8	US-10-719-900-341
6	19.4	77.6	30752	5	US-10-087-192-745
7	19.4	77.6	90100	3	US-09-997-722-295
8	19.2	76.8	626	4	US-09-925-065A-496554
9	19.2	76.8	626	4	US-09-925-065A-496555
10	19.2	76.8	1276	5	US-10-243-820-6
11	19.2	76.8	1276	5	US-10-243-817-6
12	19.2	76.8	1392	7	US-10-282-122A-24223
13	19.2	76.8	29956	3	US-09-997-722-229
14	19.2	76.8	123192	5	US-10-175-523-71
15	19.2	76.8	123192	10	US-11-099-266-71
16	18.8	75.2	43445	8	US-10-712-795-334
17	18.8	75.2	43445	8	US-10-920-612-334
18	18.8	75.2	54775	7	US-10-741-601-5659
19	18.8	75.2	54775	8	US-10-741-600-17684
20	18.6	74.4	502	7	US-10-767-701-24439
21	18.6	74.4	667	5	US-10-027-632-234119
22	18.6	74.4	667	6	US-10-027-632-234119
23	18.6	74.4	769	5	US-10-027-632-162620

c	24	18.6	74.4	769	6	US-10-027-632-162620	Sequence 162620,
	25	18.6	74.4	1305	5	US-10-027-632-204028	Sequence 204028,
	26	18.6	74.4	1305	6	US-10-027-632-204028	Sequence 204028,
c	27	18.6	74.4	7656	9	US-10-621-911A-7	Sequence 7, Appli
c	28	18.6	74.4	7656	9	US-10-646-390A-7	Sequence 7, Appli
	29	18.6	74.4	39405	5	US-10-087-192-1285	Sequence 1285, Ap
	30	18.6	74.4	39790	5	US-10-087-192-1423	Sequence 1423, Ap
c	31	18.6	74.4	47115	7	US-10-052-482-133	Sequence 133, Ap
c	32	18.6	74.4	49281	5	US-10-087-192-301	Sequence 301, App
c	33	18.6	74.4	64482	7	US-10-322-696-151	Sequence 151, App
c	34	18.6	74.4	89213	7	US-10-322-281-843	Sequence 843, App
	35	18.4	73.6	47115	7	US-10-052-482-133	Sequence 133, App
	36	18.4	72.8	600	4	US-09-925-065A-835917	Sequence 835917,
	37	17.8	71.2	331	3	US-09-728-445-172	Sequence 172, App
	38	17.8	71.2	331	3	US-10-964-549-172	Sequence 172, App
	39	17.8	71.2	4333	6	US-10-359-050-17	Sequence 2, Appli
	40	17.8	71.2	5409	9	US-10-685-837-2	Sequence 18, Appli
	41	17.8	71.2	10491	6	US-10-359-050-18	Sequence 43, Appli
	42	17.8	71.2	11176	3	US-09-738-968-43	Sequence 43, Appli
	43	17.8	71.2	11784	6	US-10-359-050-20	Sequence 106, App
	44	17.8	71.2	11784	6	US-10-014-099F-106	Sequence 12, Appli
	45	17.8	71.2	12538	6	US-10-359-050-12	Sequence 1, Appli
	46	17.8	71.2	12607	8	US-10-789-465-1	Sequence 1, Appli
	47	17.8	71.2	12607	8	US-10-867-628-1	Sequence 13, Appli
	48	17.8	71.2	12645	6	US-10-359-050-13	Sequence 1, Appli
	49	17.8	71.2	13139	9	US-10-685-837-1	Sequence 4, Appli
	50	17.8	71.2	14947	9	US-10-685-837-4	Sequence 8, Appli
	51	17.8	71.2	15174	9	US-10-685-837-8	Sequence 6, Appli
	52	17.8	71.2	15199	9	US-10-685-837-6	Sequence 171, Ap
c	53	17.8	71.2	21635	5	US-10-087-192-1171	Sequence 97, Appli
	54	17.8	71.2	22114	7	US-10-322-281-97	Sequence 805, App
	55	17.8	71.2	25543	5	US-10-087-192-805	Sequence 1777, Ap
	56	17.8	71.2	26666	8	US-10-087-192-1777	Sequence 13, Appli
	57	17.8	71.2	28446	5	US-10-331-053-13	Sequence 2035, Ap
c	58	17.8	71.2	30654	5	US-10-087-192-2035	Sequence 55, Appli
	59	17.8	71.2	32591	6	US-10-085-117-55	Sequence 445, App
	60	17.8	71.2	35183	7	US-10-322-281-445	Sequence 421, App
	61	17.8	71.2	35359	6	US-10-087-192-421	Sequence 193, App
c	62	17.8	71.2	38646	6	US-10-085-117-193	Sequence 1603, Ap
c	63	17.8	71.2	43591	5	US-10-087-192-1603	Sequence 349, App
	64	17.8	71.2	44372	5	US-10-087-192-349	Sequence 243, App
	65	17.8	71.2	47299	7	US-10-322-281-243	Sequence 13, Appli
c	66	17.8	71.2	49088	5	US-10-087-192-13	Sequence 739, App
	67	17.8	71.2	49826	5	US-10-087-192-739	Sequence 79, Appli
	68	17.8	71.2	54303	8	US-10-417-375-79	Sequence 811, App
	69	17.8	71.2	55544	5	US-10-087-192-811	Sequence 47, Appli
c	70	17.8	71.2	56773	8	US-10-331-053-47	Sequence 1975, Ap
	71	17.8	71.2	58665	5	US-10-087-192-1975	Sequence 841, App
	72	17.8	71.2	65559	5	US-10-087-192-841	Sequence 841, App
c	73	17.8	71.2	65559	5	US-10-087-192-841	Sequence 163, App
	74	17.8	71.2	65952	7	US-10-322-696-163	Sequence 423, App
c	75	17.8	71.2	66126	7	US-10-322-281-423	Sequence 1219, Ap
c	76	17.8	71.2	69515	5	US-10-087-192-1219	Sequence 217, App
	77	17.8	71.2	70215	5	US-10-087-192-217	Sequence 167, App
c	78	17.8	71.2	72530	7	US-10-322-281-265	Sequence 41, Appli
c	79	17.8	71.2	79356	7	US-10-322-281-167	Sequence 747, App
c	80	17.8	71.2	80423	7	US-10-367-094-41	Sequence 127, App
	81	17.8	71.2	84410	7	US-10-322-281-747	Sequence 703, App
	82	17.8	71.2	85435	6	US-10-085-117-127	Sequence 201, App
	83	17.8	71.2	88493	7	US-10-322-281-703	Sequence 703, App
c	84	17.8	71.2	90743	6	US-10-085-117-271	Sequence 181, App
	85	17.8	71.2	96389	7	US-10-052-482-181	Sequence 199, App
	86	17.8	71.2	96599	3	US-09-997-722-199	Sequence 53, Appli
	87	17.8	71.2	102145	8	US-10-331-053-53	Sequence 895, App
	88	17.8	71.2	111252	5	US-10-087-192-895	Sequence 1, Appli
c	89	17.8	71.2	117329	5	US-10-987-384-1	Sequence 87, Appli
	90	17.8	71.2	130427	5	US-10-175-523-87	Sequence 87, Appli
	91	17.8	71.2	130427	10	US-11-099-266-87	Sequence 283, App
c	92	17.8	71.2	155579	6	US-10-085-117-283	Sequence 139, App
	93	17.8	71.2	168976	6	US-10-085-117-139	Sequence 24, Appli
	94	17.8	71.2	171936	6	US-10-465-071-24	Sequence 24, Appli
	95	17.8	71.2	171936	6	US-10-025-966A-24	Sequence 16, Appli
c	96	17.8	71.2	215980	3	US-09-972-546-16	


```
RESULT 4
US-10-322-281-679
; Sequence 679, Application US/10322281
; Publication No. US20040126762A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc S. Malandro
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: 529452001000
; CURRENT APPLICATION NUMBER: US/10/322,281
; CURRENT FILING DATE: 2002-12-17
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 679
; LENGTH: 36599
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(36599)
; OTHER INFORMATION: n = A,T,C or G
US-10-322-281-679

Query Match      87.2%; Score 21.8; DB 7; Length 36599;
Best Local Similarity 92.0%; Pred. No. 2.9;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAAGCGCAGACTGGTCTTCC 25
Db 35597 AAAAAAGCGCAGACTGGTCTTCC 35621

RESULT 5
US-10-719-900-341
; Sequence 341, Application US/10719900
; Publication No. US20050026164A1
; GENERAL INFORMATION:
; APPLICANT: Xue Mei Zhou
; TITLE OF INVENTION: Methods of Genetic Analysis of Mouse
; FILE REFERENCE: 3528.1
; CURRENT APPLICATION NUMBER: US/10/719,900
; CURRENT FILING DATE: 2003-11-20
; PRIOR APPLICATION NUMBER: 60/427,808
; PRIOR FILING DATE: 2002-11-20
; NUMBER OF SEQ ID NOS: 982914
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 341
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-719-900-341

Query Match      83.2%; Score 20.8; DB 8; Length 25;
Best Local Similarity 91.7%; Pred. No. 4.8;
Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AAAAAAGCGCAGACTGGTCTTCC 25
Db 1 AAAAAAGCGCTCACTGGTCTTCC 24

RESULT 6
US-10-087-192-745
; Sequence 745, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01

; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 745
; LENGTH: 30752
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(30752)
; OTHER INFORMATION: n = A,T,C or G
US-10-087-192-745

Query Match      77.6%; Score 19.4; DB 5; Length 30752;
Best Local Similarity 95.2%; Pred. No. 43;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 AAGAGCGCAGACTGGTCTTCC 25
Db 3215 AAGAGCGCAGACTGGTCTTCC 3235

RESULT 7
US-09-997-722-295
; Sequence 295, Application US/09997722
; Publication No. US20040072154A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David
; APPLICANT: Engelhard, Eric
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
; FILE REFERENCE: A-71171/RMS/DCF
; CURRENT APPLICATION NUMBER: US/09/997,722
; CURRENT FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 301
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 295
; LENGTH: 90100
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1696)..(5225)
; OTHER INFORMATION: "n" at positions 1696 through 5225 can be any base.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (10303)..(10322)
; OTHER INFORMATION: "n" at positions 10303 through 10322 can be any base.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (32410)..(32429)
; OTHER INFORMATION: "n" at positions 32410 through 32429 can be any base.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (34679)..(35400)
; OTHER INFORMATION: "n" at positions 34679 through 35400 can be any base.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (38214)..(43788)
; OTHER INFORMATION: "n" at positions 38214 through 43788 can be any base.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (45408)..(46631)
; OTHER INFORMATION: "n" at positions 45408 through 46631 can be any base.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (47453)..(47810)
```

OTHER INFORMATION: "n" at positions 47453 through 47810 can be any base.
FEATURE:
NAME/KEY: misc feature
LOCATION: (49455)..(49572)
OTHER INFORMATION: "n" at positions 49495 through 49572 can be any base.
FEATURE:
NAME/KEY: misc feature
LOCATION: (52976)..(53745)
OTHER INFORMATION: "n" at positions 52976 through 53745 can be any base.
FEATURE:
NAME/KEY: misc feature
LOCATION: (59340)..(60268)
OTHER INFORMATION: "n" at positions 59340 through 60268 can be any base.
FEATURE:
NAME/KEY: misc feature
LOCATION: (76210)..(76409)
OTHER INFORMATION: "n" at positions 76210 through 76409 can be any base.
FEATURE:
NAME/KEY: misc feature
LOCATION: (83590)..(83609)
OTHER INFORMATION: "n" at positions 83590 through 83609 can be any base.
FEATURE:
NAME/KEY: misc feature
LOCATION: (88024)..(88383)
OTHER INFORMATION: "n" at positions 88024 through 88383 can be any base.
US-09-997-722-295

Query Match 77.6%; Score 19.4; DB 3; Length 90100;
Best Local Similarity 95.2%; Pred. No. 47;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 5 AAGAGCGCAGCTGCTCTCC 25
Db 50180 AAGAGCGCAGCTGCTCTCC 50200

RESULT 8
US-09-925-065A-496554
; Sequence 496554, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 496554
; LENGTH: 626
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-496554

Query Match 76.8%; Score 19.2; DB 4; Length 626;
Best Local Similarity 87.5%; Pred. No. 38;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AAAAAAGAGCGCAGCTGCTCTCC 24
Db 284 AAAAAAGAGCTCAGAAATGCTCTTC 307

RESULT 9
US-09-925-065A-496555
; Sequence 496555, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 496555
; LENGTH: 626
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-496555

Query Match 76.8%; Score 19.2; DB 4; Length 626;
Best Local Similarity 87.5%; Pred. No. 38;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AAAAAAGAGCGCAGCTGCTCTTC 24
Db 284 AAAAAAGAGTCAAGAAATGCTCTTC 307

RESULT 10
US-10-243-820-6/c
; Sequence 6, Application US/10243820
; Publication No. US20030101471A1
; GENERAL INFORMATION:
; APPLICANT: BALTIMORE, DAVID
; APPLICANT: HONG, ELIZABETH J.
; APPLICANT: LOIS-CABALLE, CARLOS
; APPLICANT: PEASE, SHIRLEY
; TITLE OF INVENTION: METHOD FOR PRODUCING TRANSGENIC ANIMALS
; FILE REFERENCE: CALTE.009A
; CURRENT APPLICATION NUMBER: US/10/243,820
; CURRENT FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 60/322,031
; PRIOR FILING DATE: 2001-09-13
; PRIOR APPLICATION NUMBER: 60/347,782
; PRIOR FILING DATE: 2002-01-09
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1276
; TYPE: DNA
; ORGANISM: Murine
US-10-243-820-6

Query Match 76.8%; Score 19.2; DB 5; Length 1276;
Best Local Similarity 87.5%; Pred. No. 41;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 AAAAAAGAGCGCAGCTGCTCTTCC 25
Db 66 AAAAAAGAGCACTGACTGCTCTTCC 43

RESULT 11
 US-10-243-817-6/c
 ; Sequence 6, Application US/10243817
 ; Publication No. US20030101472A1
 ; GENERAL INFORMATION:
 ; APPLICANT: BALTIMORE, DAVID
 ; APPLICANT: HONG, ELIZABETH J.
 ; APPLICANT: LOIS-CABALLE, CARLOS
 ; APPLICANT: PEASE, SHIRLEY
 ; TITLE OF INVENTION: METHOD FOR PRODUCING TRANSGENIC ANIMALS
 ; FILE REFERENCE: CALTE.006A
 ; CURRENT APPLICATION NUMBER: US/10/243,817
 ; CURRENT FILING DATE: 2002-09-13
 ; PRIOR APPLICATION NUMBER: 60/322,031
 ; PRIOR FILING DATE: 2001-09-13
 ; PRIOR APPLICATION NUMBER: 60/347,782
 ; PRIOR FILING DATE: 2002-01-09
 ; NUMBER OF SEQ ID NOS: 9
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 6
 ; LENGTH: 1276
 ; TYPE: DNA
 ; ORGANISM: Murine
 US-10-243-817-6

Query Match 76.8%; Score 19.2; DB 5; Length 1276;
 Best Local Similarity 87.5%; Pred. No. 41;
 Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAAAGCGCAGACTGGTCTTCC 25
 |||||
 Db 66 AAAAAAGCGCAGACTGGTCTTCC 43

RESULT 12
 US-10-282-122A-24223/c
 ; Sequence 24223, Application US/10282122A
 ; Publication No. US20040029129A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Liangsu
 ; APPLICANT: Zamudio, Carlos
 ; APPLICANT: Malone, Cheryl
 ; APPLICANT: Haselbeck, Robert
 ; APPLICANT: Ohlsen, Kari
 ; APPLICANT: Zyskind, Judith
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Trawick, John
 ; APPLICANT: Carr, Grant
 ; APPLICANT: Yamamoto, Robert
 ; APPLICANT: Forsyth, R.
 ; APPLICANT: Xu, H.
 ; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
 ; FILE REFERENCE: ELITRA.034A
 ; CURRENT APPLICATION NUMBER: US/10/282,122A
 ; CURRENT FILING DATE: 2003-02-20
 ; PRIOR APPLICATION NUMBER: 60/191,078
 ; PRIOR FILING DATE: 2000-03-21
 ; PRIOR APPLICATION NUMBER: 60/206,848
 ; PRIOR FILING DATE: 2000-05-23
 ; PRIOR APPLICATION NUMBER: 60/207,727
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: 60/230,335
 ; PRIOR FILING DATE: 2000-09-06
 ; PRIOR APPLICATION NUMBER: 60/230,347
 ; PRIOR FILING DATE: 2000-09-09
 ; PRIOR APPLICATION NUMBER: 60/242,578
 ; PRIOR FILING DATE: 2000-10-23
 ; PRIOR APPLICATION NUMBER: 60/253,625
 ; PRIOR FILING DATE: 2000-11-27
 ; PRIOR APPLICATION NUMBER: 60/257,931
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: 60/267,636
 ; PRIOR FILING DATE: 2001-02-09

; PRIOR APPLICATION NUMBER: 60/269,308
 ; PRIOR FILING DATE: 2001-02-16
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 78614
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 24223
 ; LENGTH: 1392
 ; TYPE: DNA
 ; ORGANISM: Listeria monocytogenes
 US-10-282-122A-24223

Query Match 76.8%; Score 19.2; DB 7; Length 1392;
 Best Local Similarity 87.5%; Pred. No. 41;
 Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAAGCGCAGACTGGTCTTCC 24
 |||||
 Db 132 AAAAAAGCGCGACTGATTTTC 109

RESULT 13
 US-09-997-722-229/c
 ; Sequence 229, Application US/09997722
 ; Publication No. US20040072154A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Morris, David
 ; APPLICANT: Engelhard, Eric
 ; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
 ; FILE REFERENCE: A-71171/RMS/DCF
 ; CURRENT APPLICATION NUMBER: US/09/997,722
 ; CURRENT FILING DATE: 2001-11-30
 ; PRIOR APPLICATION NUMBER: US 09/747,377
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: US 09/798,586
 ; PRIOR FILING DATE: 2001-03-02
 ; NUMBER OF SEQ ID NOS: 301
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 229
 ; LENGTH: 29956
 ; TYPE: DNA
 ; ORGANISM: Mus musculus
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (1335)..(1858)
 ; OTHER INFORMATION: "n" at positions 1335 through 1858 can be any base.
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (7091)..(7110)
 ; OTHER INFORMATION: "n" at positions 7091 through 7110 can be any base.
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (20085)..(20586)
 ; OTHER INFORMATION: "n" at positions 20085 through 20586 can be any base.
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (24372)..(24840)
 ; OTHER INFORMATION: "n" at positions 24372 through 24840 can be any base.
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (29121)..(29589)
 ; OTHER INFORMATION: "n" at positions 29121 through 29589 can be any base.
 US-09-997-722-229

Query Match 76.8%; Score 19.2; DB 3; Length 29956;
 Best Local Similarity 87.5%; Pred. No. 54;
 Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAAAAAGCGCAGACTGGTCTTCC 25
 |||||
 Db 8863 AAAAAAGCGCAGACTGGTCTTCC 8840

RESULT 14

```

US-10-175-523-71
; Sequence 71, Application US/10175523
; Publication No. US20030096264A1
; GENERAL INFORMATION:
; APPLICANT: Brockman, Jeffrey
; APPLICANT: Evans, David
; APPLICANT: Hook, Derek
; APPLICANT: Klimczak, Leszek
; APPLICANT: Laeng, Pascal
; APPLICANT: Palfreyman, Michael
; APPLICANT: Rajan, Prihi
; TITLE OF INVENTION: MULTI-PARAMETER HIGH THROUGHPUT SCREENING ASSAYS (MPHTS)
; FILE REFERENCE: 3235/11795-US3
; CURRENT APPLICATION NUMBER: US/10/175,523
; CURRENT FILING DATE: 2002-06-18
; PRIOR APPLICATION NUMBER: US 60/299,151
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: US 60/317,828
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: US 60/325,150
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/333,047
; PRIOR FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: US 60/349,936
; PRIOR FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: US 60/361,834
; PRIOR FILING DATE: 2002-03-04
; NUMBER OF SEQ ID NOS: 197
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 71
; LENGTH: 123192
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-175-523-71

Query Match 76.8%; Score 19.2; DB 5; Length 123192;
Best Local Similarity 87.5%; Pred. No. 61;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAAGCGCAGACTGCTCTTC 24
Db 73891 AAAAAAGCGCAGACTGCTCTTC 73914

RESULT 15
US-11-099-266-71
; Sequence 71, Application US/11099266
; Publication No. US20050181433A1
; GENERAL INFORMATION:
; APPLICANT: Brockman, Jeffrey
; APPLICANT: Evans, David
; APPLICANT: Hook, Derek
; APPLICANT: Klimczak, Leszek
; APPLICANT: Laeng, Pascal
; APPLICANT: Palfreyman, Michael
; APPLICANT: Rajan, Prihi
; TITLE OF INVENTION: MULTI-PARAMETER HIGH THROUGHPUT SCREENING ASSAYS (MPHTS)
; FILE REFERENCE: 03235/100795-US4
; CURRENT APPLICATION NUMBER: US/11/099,266
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: US 10/175,523
; PRIOR FILING DATE: 2002-06-18
; PRIOR APPLICATION NUMBER: US 60/299,151
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: US 60/317,828
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: US 60/325,150
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/333,047
; PRIOR FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: US 60/349,936
; PRIOR FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: US 60/361,834

```

```

; PRIOR FILING DATE: 2002-03-04
; NUMBER OF SEQ ID NOS: 197
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 71
; LENGTH: 123192
; TYPE: DNA
; ORGANISM: Mus musculus
US-11-099-266-71

Query Match 76.8%; Score 19.2; DB 10; Length 123192;
Best Local Similarity 87.5%; Pred. No. 61;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAAAAGCGCAGACTGCTCTTC 24
Db 73891 AAAAAAGCGCAGACTGCTCTTC 73914

RESULT 16
US-10-712-795-334
; Sequence 334, Application US/10712795
; Publication No. US20040214325A1
; GENERAL INFORMATION:
; APPLICANT: Crooke et al.
; TITLE OF INVENTION: ANTISENSE MODULATION OF APOLIPOPROTEIN B EXPRESSION
; FILE REFERENCE: 30566/39662
; CURRENT APPLICATION NUMBER: US/10/712,795
; CURRENT FILING DATE: 2003-11-13
; PRIOR APPLICATION NUMBER: US 60/426,234
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: PCT/US03/15493
; PRIOR FILING DATE: 2003-05-13
; NUMBER OF SEQ ID NOS: 892
; SEQ ID NO 334
; LENGTH: 43445
; TYPE: DNA
; ORGANISM: H. sapiens
US-10-712-795-334

Query Match 75.2%; Score 18.8; DB 8; Length 43445;
Best Local Similarity 90.9%; Pred. No. 87;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAAGCGCAGACTGCTCT 22
Db 25768 AAAACAGACGTAGACTGCTCT 25789

RESULT 17
US-10-920-612-334
; Sequence 334, Application US/10920612
; Publication No. US20050009088A1
; GENERAL INFORMATION:
; APPLICANT: Crooke et al.
; TITLE OF INVENTION: ANTISENSE MODULATION OF APOLIPOPROTEIN B EXPRESSION
; FILE REFERENCE: 30566/39634A
; CURRENT APPLICATION NUMBER: US/10/920,612
; CURRENT FILING DATE: 2004-08-17
; PRIOR APPLICATION NUMBER: PCT/US03/15493
; PRIOR FILING DATE: 2003-11-15
; PRIOR APPLICATION NUMBER: US 10/712,795
; PRIOR FILING DATE: 2003-11-13
; PRIOR APPLICATION NUMBER: US 60/426,234
; PRIOR FILING DATE: 2002-11-13
; NUMBER OF SEQ ID NOS: 892
; SEQ ID NO 334
; LENGTH: 43445
; TYPE: DNA
; ORGANISM: H. sapiens
US-10-920-612-334

Query Match 75.2%; Score 18.8; DB 8; Length 43445;
Best Local Similarity 90.9%; Pred. No. 87;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AAAAAAGCGCAGACTGGTCT 22
Db 25768 AAAAAAGCGCAGACTGGTCT 25789

RESULT 18

US-10-741-601-5659
; Sequence 5659, Application US/10741601
; Publication No. US20040166519A1

; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: CL001500
; CURRENT APPLICATION NUMBER: US/10/741,601
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 26415
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 5659
; LENGTH: 54775

; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-741-601-5659

Query Match 75.2%; Score 18.8; DB 7; Length 54775;
Best Local Similarity 90.9%; Pred. No. 89;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AAAAAAGCGCAGACTGGTCT 22
Db 31373 AAAAAAGCGTAGACTGGTCT 31394

RESULT 19

US-10-741-600-17684
; Sequence 17684, Application US/10741600
; Publication No. US20050026169A1

; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 17684
; LENGTH: 54775

; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-741-600-17684

Query Match 75.2%; Score 18.8; DB 8; Length 54775;
Best Local Similarity 90.9%; Pred. No. 89;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AAAAAAGCGCAGACTGGTCT 22
Db 31373 AAAAAAGCGTAGACTGGTCT 31394

RESULT 20

US-10-767-701-24439/c
; Sequence 24439, Application US/10767701
; Publication No. US20040172684A1

; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement

; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 24439
; LENGTH: 502

; TYPE: DNA
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: 30935687
US-10-767-701-24439

Query Match 74.4%; Score 18.6; DB 7; Length 502;
Best Local Similarity 84.0%; Pred. No. 74;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 AAAAAAGCGCAGACTGGTCTTCC 25
Db 265 AAAAAAGACAGGCTCGTCTTCC 241

RESULT 21

US-10-027-632-234119
; Sequence 234119, Application US/10027632
; Publication No. US20020198371A1

; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome

; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12

; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 234119
; LENGTH: 667

; TYPE: DNA
; ORGANISM: Human
US-10-027-632-234119

Query Match 74.4%; Score 18.6; DB 5; Length 667;
Best Local Similarity 84.0%; Pred. No. 76;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 AAAAAAGCGCAGACTGGTCTTCC 25
Db 600 AAAAAAGGTCAGACTTGGCTTCC 624

RESULT 22

US-10-027-632-234119
; Sequence 234119, Application US/10027632
; Publication No. US20030204075A9

; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632

```
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 234119
; LENGTH: 667
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-234119
```

```
Query Match 74.4%; Score 18.6; DB 6; Length 667;
Best Local Similarity 84.0%; Pred. No. 76;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
```

```
QY 1 AAAAAAGCGCAGACTGGCTTCC 25
|||||
DB 600 AAAAAAGGTCAGACTTGGCTTCC 624
```

```
RESULT 23
US-10-027-632-162620/c
; Sequence 162620, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 162620
; LENGTH: 769
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-162620
```

```
Query Match 74.4%; Score 18.6; DB 5; Length 769;
Best Local Similarity 84.0%; Pred. No. 77;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
```

```
QY 1 AAAAAAGCGCAGACTGGCTTCC 25
|||||
DB 171 AAAAAAGGTCAGACTTGGCTTCC 147
```

```
RESULT 24
US-10-027-632-162620/c
; Sequence 162620, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 162620
; LENGTH: 769
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-162620
```

```
Query Match 74.4%; Score 18.6; DB 6; Length 769;
Best Local Similarity 84.0%; Pred. No. 77;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
```

```
QY 1 AAAAAAGCGCAGACTGGCTTCC 25
|||||
DB 171 AAAAAAGGTCAGACTTGGCTTCC 147
```

```
RESULT 25
US-10-027-632-204028
; Sequence 204028, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 204028
; LENGTH: 1305
; TYPE: DNA
```

```

; ORGANISM: Human
US-10-027-632-204028

Query Match          74.4%; Score 18.6; DB 5; Length 1305;
Best Local Similarity 84.0%; Pred. No. 80;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAAGCGCAGACTGGTCTTCC 25
    ||||| ||||| ||||| ||||| |||||
Db 518 AAAAAAGGTCAGACTTGGCTTCC 542

RESULT 26
US-10-027-632-204028
; Sequence 204028, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; POLYMORPHISMS IN THE HUMAN GENOME
; FILE REFERENCE: 108927.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 204028
; LENGTH: 1305
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-204028

Query Match          74.4%; Score 18.6; DB 6; Length 1305;
Best Local Similarity 84.0%; Pred. No. 80;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAAGCGCAGACTGGTCTTCC 25
    ||||| ||||| ||||| ||||| |||||
Db 518 AAAAAAGGTCAGACTTGGCTTCC 542

RESULT 27
US-10-621-911A-7/c
; Sequence 7, Application US/10621911A
; Publication No. US20050054823A1
; GENERAL INFORMATION:
; APPLICANT: SAITOU, Mitinori
; TITLE OF INVENTION: Genes
; FILE REFERENCE: 674558-2002
; CURRENT APPLICATION NUMBER: US/10/621,911A
; CURRENT FILING DATE: 2003-07-17
; PRIOR APPLICATION NUMBER: PCT/GB02/00215
; PRIOR FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: GB 0101300.2
; PRIOR FILING DATE: 2001-01-18
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: SeqWin99, version 1.02
; SEQ ID NO 7
; LENGTH: 7656
; TYPE: DNA
; ORGANISM: Rattus sp
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (7471)..(7471)
; OTHER INFORMATION: "n" is an unknown nucleotide
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (7554)..(7554)
; OTHER INFORMATION: "n" is an unknown nucleotide
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (7608)..(7608)
; OTHER INFORMATION: "n" is an unknown nucleotide
US-10-646-390A-7/c
; Sequence 7, Application US/10646390A
; Publication No. US20050054824A1
; GENERAL INFORMATION:
; APPLICANT: SAITOU, Mitinori
; TITLE OF INVENTION: Genes
; FILE REFERENCE: 674558-2002.1
; CURRENT APPLICATION NUMBER: US/10/646,390A
; CURRENT FILING DATE: 2003-08-21
; PRIOR APPLICATION NUMBER: 10/621,911
; PRIOR FILING DATE: 2003-07-17
; PRIOR APPLICATION NUMBER: PCT/GB02/00215
; PRIOR FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: GB 0101300.2
; PRIOR FILING DATE: 2001-01-18
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: SeqWin99, version 1.02
; SEQ ID NO 7
; LENGTH: 7656
; TYPE: DNA
; ORGANISM: Rattus sp
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (7471)..(7471)
; OTHER INFORMATION: "n" is an unknown nucleotide
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (7554)..(7554)
; OTHER INFORMATION: "n" is an unknown nucleotide
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (7608)..(7608)
; OTHER INFORMATION: "n" is an unknown nucleotide
US-10-646-390A-7

Query Match          74.4%; Score 18.6; DB 9; Length 7656;
Best Local Similarity 84.0%; Pred. No. 94;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAAGCGCAGACTGGTCTTCC 25
    ||||| ||||| ||||| ||||| |||||
Db 2113 AAAAAAGGTCGACTGCTCTTCC 2089

RESULT 28
US-10-646-390A-7/c
; Sequence 7, Application US/10646390A
; Publication No. US20050054824A1
; GENERAL INFORMATION:
; APPLICANT: SAITOU, Mitinori
; TITLE OF INVENTION: Genes
; FILE REFERENCE: 674558-2002.1
; CURRENT APPLICATION NUMBER: US/10/646,390A
; CURRENT FILING DATE: 2003-08-21
; PRIOR APPLICATION NUMBER: 10/621,911
; PRIOR FILING DATE: 2003-07-17
; PRIOR APPLICATION NUMBER: PCT/GB02/00215
; PRIOR FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: GB 0101300.2
; PRIOR FILING DATE: 2001-01-18
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: SeqWin99, version 1.02
; SEQ ID NO 7
; LENGTH: 7656
; TYPE: DNA
; ORGANISM: Rattus sp
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (7471)..(7471)
; OTHER INFORMATION: "n" is an unknown nucleotide
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (7554)..(7554)
; OTHER INFORMATION: "n" is an unknown nucleotide
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (7608)..(7608)
; OTHER INFORMATION: "n" is an unknown nucleotide
US-10-646-390A-7

Query Match          74.4%; Score 18.6; DB 9; Length 7656;
Best Local Similarity 84.0%; Pred. No. 94;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAAGCGCAGACTGGTCTTCC 25
    ||||| ||||| ||||| ||||| |||||
Db 2113 AAAAAAGGTCGACTGCTCTTCC 2089

RESULT 29
US-10-087-192-1285
```

```
/ Sequence 1285, Application US/10087192
/ Publication No. US20020182586A1
/ GENERAL INFORMATION:
/ APPLICANT: Engelhard, David W.
/ APPLICANT: Engelhard, Eric K.
/ TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
/ TITLE OF INVENTION: CANCER
/ FILE REFERENCE: 529452000122
/ CURRENT APPLICATION NUMBER: US/10/087,192
/ CURRENT FILING DATE: 2002-03-01
/ PRIOR APPLICATION NUMBER: US 09/747,377
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: US 09/798,586
/ PRIOR FILING DATE: 2001-03-02
/ NUMBER OF SEQ ID NOS: 2059
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 1285
/ LENGTH: 39405
/ TYPE: DNA
/ ORGANISM: Mus musculus
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1)...(39405)
/ OTHER INFORMATION: n = A,T,C or G
US-10-087-192-1285
```

```
Query Match 74.4%; Score 18.6; DB 5; Length 39405;
Best Local Similarity 84.0%; Pred. No. 1.1e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
```

```
QY 1 AAAAAAGAGCGAGACTGCTCTTCC 25
Db 9488 AGAAAGAGCATGACTGCTCTTCC 9512
```

RESULT 30

```
US-10-087-192-1423
/ Sequence 1423, Application US/10087192
/ Publication No. US20020182586A1
/ GENERAL INFORMATION:
/ APPLICANT: Engelhard, David W.
/ APPLICANT: Engelhard, Eric K.
/ TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
/ TITLE OF INVENTION: CANCER
/ FILE REFERENCE: 529452000122
/ CURRENT APPLICATION NUMBER: US/10/087,192
/ CURRENT FILING DATE: 2002-03-01
/ PRIOR APPLICATION NUMBER: US 09/747,377
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: US 09/798,586
/ PRIOR FILING DATE: 2001-03-02
/ NUMBER OF SEQ ID NOS: 2059
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 1423
/ LENGTH: 39790
/ TYPE: DNA
/ ORGANISM: Mus musculus
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1)...(39790)
/ OTHER INFORMATION: n = A,T,C or G
US-10-087-192-1423
```

```
Query Match 74.4%; Score 18.6; DB 5; Length 39790;
Best Local Similarity 84.0%; Pred. No. 1.1e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
```

```
QY 1 AAAAAAGAGCGAGACTGCTCTTCC 25
Db 15595 AATAAGAGCACTGACTGCTCTTCC 15619
```

RESULT 31

```
US-10-052-482-133/c
/ Sequence 133, Application US/10052482
/ Publication No. US20040072264A1
/ GENERAL INFORMATION:
/ APPLICANT: Engelhard, Eric
/ APPLICANT: Morris, David
/ TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
/ FILE REFERENCE: A-71087/RMS/DCF
/ CURRENT APPLICATION NUMBER: US/10/052,482
/ CURRENT FILING DATE: 2002-08-15
/ PRIOR APPLICATION NUMBER: US 09/747,377
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: US 09/798,586
/ PRIOR FILING DATE: 2001-03-02
/ NUMBER OF SEQ ID NOS: 241
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 133
/ LENGTH: 47115
/ TYPE: DNA
/ ORGANISM: Mus musculus
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1171)..(1270)
/ OTHER INFORMATION: "n" at positions 1171 to 1270 can be any base
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (28066)..(28402)
/ OTHER INFORMATION: "n" at positions 28066 to 28402 can be any base
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (34855)..(34891)
/ OTHER INFORMATION: "n" at positions 34855 to 34891 can be any base
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (37116)..(37389)
/ OTHER INFORMATION: "n" at positions 37116 to 37389 can be any base
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (43130)..(44280)
/ OTHER INFORMATION: "n" at positions 43130 to 44280 can be any base
US-10-052-482-133
```

```
Query Match 74.4%; Score 18.6; DB 7; Length 47115;
Best Local Similarity 84.0%; Pred. No. 1.1e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
```

```
QY 1 AAAAAAGAGCGAGACTGCTCTTCC 25
Db 5679 AAAAAAGAGCATGACTGCTCTTCC 5655
```

RESULT 32

```
US-10-087-192-301/c
/ Sequence 301, Application US/10087192
/ Publication No. US20020182586A1
/ GENERAL INFORMATION:
/ APPLICANT: Morris, David W.
/ APPLICANT: Engelhard, Eric K.
/ TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
/ TITLE OF INVENTION: CANCER
/ FILE REFERENCE: 529452000122
/ CURRENT APPLICATION NUMBER: US/10/087,192
/ CURRENT FILING DATE: 2002-03-01
/ PRIOR APPLICATION NUMBER: US 09/747,377
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: US 09/798,586
/ PRIOR FILING DATE: 2001-03-02
/ NUMBER OF SEQ ID NOS: 2059
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 301
/ LENGTH: 49281
/ TYPE: DNA
/ ORGANISM: Mus musculus
```

FEATURE:
NAME/KEY: misc feature
LOCATION: (1)...(49281)
OTHER INFORMATION: n = A,T,C or G
US-10-087-192-301

Query Match 74.4%; Score 18.6; DB 5; Length 49281;
Best Local Similarity 84.0%; Pred. No. 1.1e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAAGCGCAGACTGCTCTTC 25
Db 81877 AAAAAAGAGTGTGCGACTGCTCTTC 81853

RESULT 35
US-10-052-482-133
Sequence 133, Application US/10052482
Publication No. US20040072264A1
GENERAL INFORMATION:
APPLICANT: Engelhard, Eric
TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
FILE REFERENCE: A-71087/RMS/DCF
CURRENT APPLICATION NUMBER: US/10/052,482
CURRENT FILING DATE: 2002-08-15
PRIOR APPLICATION NUMBER: US 09/747,377
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: US 09/798,586
PRIOR FILING DATE: 2001-03-02
NUMBER OF SEQ ID NOS: 241
SOFTWARE: PatentIn version 3.1
SEQ ID NO 133
LENGTH: 47115
TYPE: DNA
ORGANISM: Mus musculus
FEATURE:
NAME/KEY: misc feature
LOCATION: (1171)..(1270)
OTHER INFORMATION: "n" at positions 1171 to 1270 can be any base
FEATURE:
NAME/KEY: misc feature
LOCATION: (28056)..(28402)
OTHER INFORMATION: "n" at positions 28066 to 28402 can be any base
FEATURE:
NAME/KEY: misc feature
LOCATION: (34855)..(34891)
OTHER INFORMATION: "n" at positions 34855 to 34891 can be any base
FEATURE:
NAME/KEY: misc feature
LOCATION: (37116)..(37389)
OTHER INFORMATION: "n" at positions 37116 to 37389 can be any base
FEATURE:
NAME/KEY: misc feature
LOCATION: (43130)..(44280)
OTHER INFORMATION: "n" at positions 43130 to 44280 can be any base
US-10-052-482-133

Query Match 73.6%; Score 18.4; DB 7; Length 47115;
Best Local Similarity 95.0%; Pred. No. 1.4e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 AAGAGCGCAGACTGCTCTTC 24
Db 34115 AAGAGCGCAGACTGCTCTTC 34134

RESULT 36
US-09-925-065A-835917
Sequence 835917, Application US/09925065A
Publication No. US20050228172A9
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single
Nucleotide Polymorphisms in the Human Genome
FILE REFERENCE: 108827.135
CURRENT APPLICATION NUMBER: US/09/925,065A
CURRENT FILING DATE: 2001-08-08
PRIOR APPLICATION NUMBER: US 60/243,096
PRIOR FILING DATE: 2000-10-24
PRIOR APPLICATION NUMBER: US 60/252,147
PRIOR FILING DATE: 2000-11-20

FEATURE:
NAME/KEY: misc feature
LOCATION: (1)...(49281)
OTHER INFORMATION: n = A,T,C or G
US-10-087-192-301

Query Match 74.4%; Score 18.6; DB 5; Length 49281;
Best Local Similarity 84.0%; Pred. No. 1.1e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAAGCGCAGACTGCTCTTC 25
Db 11432 AAAAAAGAGTGTGCGACTGCTCTTC 11408

RESULT 33
US-10-322-696-151/c
Sequence 151, Application US/10322696
Publication No. US20040166490A1
GENERAL INFORMATION:
APPLICANT: Morris, David W.
TITLE OF INVENTION: NOVEL THERAPEUTIC TARGETS IN CANCER
FILE REFERENCE: 529452001200
CURRENT APPLICATION NUMBER: US/10/322,696
CURRENT FILING DATE: 2003-10-17
NUMBER OF SEQ ID NOS: 186
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 151
LENGTH: 64482
TYPE: DNA
ORGANISM: Mus musculus
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)...(64482)
OTHER INFORMATION: n = A,T,C or G
US-10-322-696-151

Query Match 74.4%; Score 18.6; DB 7; Length 64482;
Best Local Similarity 84.0%; Pred. No. 1.1e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAAGCGCAGACTGCTCTTC 25
Db 53861 AATAAGAGCAGTGTCTCTTC 53837

RESULT 34
US-10-322-281-843/c
Sequence 843, Application US/10322281
Publication No. US20040126762A1
GENERAL INFORMATION:
APPLICANT: David W. Morris
TITLE OF INVENTION: Novel Compositions and Methods in Cancer
FILE REFERENCE: 529452001000
CURRENT APPLICATION NUMBER: US/10/322,281
CURRENT FILING DATE: 2002-12-17
NUMBER OF SEQ ID NOS: 866
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 843
LENGTH: 89213
TYPE: DNA
ORGANISM: Mus musculus
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)...(89213)
OTHER INFORMATION: n = A,T,C or G
US-10-322-281-843

Query Match 74.4%; Score 18.6; DB 7; Length 89213;
Best Local Similarity 84.0%; Pred. No. 1.2e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 835917
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-835917

Query Match 72.8%; Score 18.2; DB 4; Length 600;
Best Local Similarity 87.0%; Pred. No. 1.2e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AAGAGCGCAGACTGGTCTTC 24
DB 334 AAGAGCGGTAGACTGGTATTC 356

RESULT 37
US-09-728-445-172
; Sequence 172, Application US/09728445
; Patent No. US20020102543A1
; GENERAL INFORMATION:
; APPLICANT: Friedrich, Glenn
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. US20020102543A1el Mutated Mammalian Cells and
; FILE REFERENCE: LEX-0102-USA
; CURRENT APPLICATION NUMBER: US/09/728,445
; CURRENT FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/168,358
; PRIOR FILING DATE: 1999-12-01
; NUMBER OF SEQ ID NOS: 891
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 172
; LENGTH: 331
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(331)
; OTHER INFORMATION: n = A,T,C or G
US-09-728-445-172

Query Match 71.2%; Score 17.8; DB 3; Length 331;
Best Local Similarity 90.5%; Pred. No. 1.7e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 AAGAGCGCAGACTGGTCTTC 25
DB 184 AAGAGCACAGACTGCTCTTC 204

RESULT 38
US-10-964-549-172
; Sequence 172, Application US/10964549
; Publication No. US20050186677A1
; GENERAL INFORMATION:
; APPLICANT: Friedrich, Glenn
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: Novel Mutated Mammalian Cells and Animals
; FILE REFERENCE: LEX-0286-USA
; CURRENT APPLICATION NUMBER: US/10/964,549
; CURRENT FILING DATE: 2004-10-13
; PRIOR APPLICATION NUMBER: US/09/750,456
; PRIOR FILING DATE: 2000-12-28

; PRIOR APPLICATION NUMBER: US 09/728,445
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/168,358
; PRIOR FILING DATE: 1999-12-01
; NUMBER OF SEQ ID NOS: 891
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 172
; LENGTH: 331
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(331)
; OTHER INFORMATION: n = A,T,C or G
US-10-964-549-172

Query Match 71.2%; Score 17.8; DB 9; Length 331;
Best Local Similarity 90.5%; Pred. No. 1.7e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 AAGAGCGCAGACTGGTCTTC 25
DB 184 AAGAGCACAGACTGCTCTTC 204

RESULT 39
US-10-359-050-17
; Sequence 17, Application US/10359050
; Publication No. US20030186291A1
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS DEUTSCHLAND GMBH
; TITLE OF INVENTION: GENETICALLY ENGINEERED PHIC31-INTEGRASE GENES
; FILE REFERENCE: AR03-001
; CURRENT APPLICATION NUMBER: US/10/359,050
; CURRENT FILING DATE: 2003-02-05
; PRIOR APPLICATION NUMBER: US 60/354,741
; PRIOR FILING DATE: 2002-02-06
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 17
; LENGTH: 4333
; TYPE: DNA
; ORGANISM: ROSA 3'
US-10-359-050-17

Query Match 71.2%; Score 17.8; DB 6; Length 4333;
Best Local Similarity 90.5%; Pred. No. 2.2e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 AAGAGCGCAGACTGGTCTTC 25
DB 1308 AAGAGCACAGACTGCTCTTC 1328

RESULT 40
US-10-685-837-2
; Sequence 2, Application US/10685837
; Publication No. US20050071893A1
; GENERAL INFORMATION:
; APPLICANT: Seibler, Jost
; APPLICANT: Schwenk, Frieder
; APPLICANT: Kuhn, Ralf
; APPLICANT: Kuter-Luks, Birgit
; TITLE OF INVENTION: siRNA mediated gene silencing in transgenic animals
; FILE REFERENCE: 022698us JH/BM
; CURRENT APPLICATION NUMBER: US/10/685,837
; CURRENT FILING DATE: 2003-10-15
; PRIOR APPLICATION NUMBER: US60/420,476
; PRIOR FILING DATE: 2002-10-22
; PRIOR APPLICATION NUMBER: US60/467,814
; PRIOR FILING DATE: 2003-05-03
; PRIOR APPLICATION NUMBER: US60/485,969
; PRIOR FILING DATE: 2003-07-03


```
; PRIOR APPLICATION NUMBER: 02023283.1
; PRIOR FILING DATE: 2002-10-17
; NUMBER OF SEQ ID NOS: 220
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 5409
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: sequence of
; OTHER INFORMATION: homology region
; US-10-685-837-2

Query Match          71.2%; Score 17.8; DB 9; Length 5409;
Best Local Similarity 90.5%; Pred. No. 2.2e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      5 AAGAGCGCAGACTGGTCTTCC 25
      ||||| ||||| ||||| |||||
Db      2384 AAGAGCACAGACTGCTCTTCC 2404

RESULT 41
US-10-359-050-18
; Sequence 18, Application US/10359050
; Publication No. US20030186291A1
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS DEUTSCHLAND GMBH
; TITLE OF INVENTION: GENETICALLY ENGINEERED PHIC31-INTEGRASE GENES
; FILE REFERENCE: AR03-001
; CURRENT APPLICATION NUMBER: US/10/359,050
; CURRENT FILING DATE: 2003-02-05
; PRIOR APPLICATION NUMBER: US 60/354,741
; PRIOR FILING DATE: 2002-02-06
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 18
; LENGTH: 10491
; TYPE: DNA
; ORGANISM: PROSA 12
; US-10-359-050-18

Query Match          71.2%; Score 17.8; DB 6; Length 10491;
Best Local Similarity 90.5%; Pred. No. 2.4e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      5 AAGAGCGCAGACTGGTCTTCC 25
      ||||| ||||| ||||| |||||
Db      4954 AAGAGCACAGACTGCTCTTCC 4974

RESULT 42
US-09-738-968-43
; Sequence 43, Application US/09738968
; Patent No. US20010037016A1
; GENERAL INFORMATION:
; APPLICANT: Contag, Pamela R.
; APPLICANT: Purchio, Anthony
; APPLICANT: Zhang, Ning
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR SCREENING FOR ANGIOGENESIS
; TITLE OF INVENTION: MODULATING COMPOUNDS
; FILE REFERENCE: 9400-0012.20
; CURRENT APPLICATION NUMBER: US/09/738,968
; CURRENT FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: 09/465,978
; PRIOR FILING DATE: 1999-12-16
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 43
; LENGTH: 11176
; TYPE: DNA
; ORGANISM: Mus sp.
; US-09-738-968-43
```

```
Query Match          71.2%; Score 17.8; DB 3; Length 11176;
Best Local Similarity 90.5%; Pred. No. 2.4e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY      5 AAGAGCGCAGACTGGTCTTCC 25
      ||||| ||||| ||||| |||||
Db      10526 AAGAGCACAGACTGCTCTTCC 10546
```

```
RESULT 43
US-10-359-050-20
; Sequence 20, Application US/10359050
; Publication No. US20030186291A1
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS DEUTSCHLAND GMBH
; TITLE OF INVENTION: GENETICALLY ENGINEERED PHIC31-INTEGRASE GENES
; FILE REFERENCE: AR03-001
; CURRENT APPLICATION NUMBER: US/10/359,050
; CURRENT FILING DATE: 2003-02-05
; PRIOR APPLICATION NUMBER: US 60/354,741
; PRIOR FILING DATE: 2002-02-06
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 20
; LENGTH: 11784
; TYPE: DNA
; ORGANISM: C31 substrate reporter
; US-10-359-050-20
```

```
Query Match          71.2%; Score 17.8; DB 6; Length 11784;
Best Local Similarity 90.5%; Pred. No. 2.4e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY      5 AAGAGCGCAGACTGGTCTTCC 25
      ||||| ||||| ||||| |||||
Db      8760 AAGAGCACAGACTGCTCTTCC 8780
```

```
RESULT 44
US-10-014-099F-106
; Sequence 106, Application US/10014099F
; Publication No. US20040003420A1
; GENERAL INFORMATION:
; APPLICANT: KUEHN, Ralf
; APPLICANT: FELDER, Susanne
; APPLICANT: SCHWENK, Frieder
; APPLICANT: KUESTER-LUKS, Birgit
; APPLICANT: FAUST, Nicole
; TITLE OF INVENTION: Modified Recombinase
; FILE REFERENCE: 012787wo/JH/ml
; CURRENT APPLICATION NUMBER: US/10/014,099F
; CURRENT FILING DATE: 2001-11-12
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 106
; LENGTH: 11784
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: modified
; OTHER INFORMATION: ROSA26 locus
; US-10-014-099F-106
```

```
Query Match          71.2%; Score 17.8; DB 6; Length 11784;
Best Local Similarity 90.5%; Pred. No. 2.4e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY      5 AAGAGCGCAGACTGGTCTTCC 25
      ||||| ||||| ||||| |||||
Db      8760 AAGAGCACAGACTGCTCTTCC 8780
```

RESULT 45

US-10-359-050-12
; Sequence 12, Application US/10359050
; Publication No. US20030186291A1
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS DEUTSCHLAND GMBH
; TITLE OF INVENTION: GENETICALLY ENGINEERED PHIC31-INTEGRASE GENES
; FILE REFERENCE: AR03-001
; CURRENT APPLICATION NUMBER: US/10/359,050
; CURRENT FILING DATE: 2003-02-05
; PRIOR APPLICATION NUMBER: US 60/354,741
; PRIOR FILING DATE: 2002-02-06
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 12
; LENGTH: 12538
; TYPE: DNA
; ORGANISM: pROSA-SA-C31-Int (CNLS)
US-10-359-050-12

Query Match 71.2%; Score 17.8; DB 6; Length 12538;
Best Local Similarity 90.5%; Pred. No. 2.4e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 AAGAGCGCAGACTGGTCTTCC 25
||||| ||||| ||||| |||||
Db 7001 AAGAGCACAGACTGCTCTTCC 7021

RESULT 46

US-10-789-465-1
; Sequence 1, Application US/10789465
; Publication No. US20040197317A1
; GENERAL INFORMATION:
; APPLICANT: Rao, Mahendra S
; TITLE OF INVENTION: Persistent Expression of Candidate Molecule in Proliferating Stem
; TITLE OF INVENTION: and Progenitor Cells for Delivery of Therapeutic Products
; FILE REFERENCE: 2923-5456.1PC
; CURRENT APPLICATION NUMBER: US/10/789,465
; CURRENT FILING DATE: 2004-02-27
; PRIOR APPLICATION NUMBER: US 60/440,152
; PRIOR FILING DATE: 2003-01-13
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 12607
; TYPE: DNA
; ORGANISM: Mus musculus, synthetic
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(2166)
; OTHER INFORMATION: Region: pUC vector
; FEATURE:
; NAME/KEY: gene
; LOCATION: (2191)..(6098)
; OTHER INFORMATION: Polr2a
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (2291)..(2477)
; OTHER INFORMATION: Exon 26 of polr2a gene
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (2817)..(2966)
; OTHER INFORMATION: Exon 27 of polr2a gene
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (3050)..(4206)
; OTHER INFORMATION: Exon 28 of polr2a gene
; FEATURE:
; NAME/KEY: 3'UTR
; LOCATION: (4207)..(6098)
; OTHER INFORMATION: gene=Polr2a

; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4469)..(5056)
; OTHER INFORMATION: Region: Internal Ribosome Entry Site (IRES)
; FEATURE:
; NAME/KEY: gene
; LOCATION: (5057)..(5860)
; OTHER INFORMATION: gene=neo
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (12140)..(12607)
; OTHER INFORMATION: Region: pUC vector
US-10-789-465-1

Query Match 71.2%; Score 17.8; DB 8; Length 12607;
Best Local Similarity 90.5%; Pred. No. 2.4e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 AAGAGCGCAGACTGGTCTTCC 25
||||| ||||| ||||| |||||
Db 11670 AAGAGCGCGACTGCTCTTCC 11690

RESULT 47

US-10-867-628-1
; Sequence 1, Application US/10867628
; Publication No. US20050048041A1
; GENERAL INFORMATION:
; APPLICANT: Rao, Mahendra S
; TITLE OF INVENTION: Persistent Expression of Candidate Molecule in Proliferating Stem
; TITLE OF INVENTION: and Progenitor Cells for Delivery of Therapeutic Products
; FILE REFERENCE: 2923-5456.1PC
; CURRENT APPLICATION NUMBER: US/10/867,628
; CURRENT FILING DATE: 2004-06-15
; PRIOR APPLICATION NUMBER: US 60/440,152
; PRIOR FILING DATE: 2003-01-13
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 12607
; TYPE: DNA
; ORGANISM: Mus musculus, synthetic
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(2166)
; OTHER INFORMATION: Region: pUC vector
; FEATURE:
; NAME/KEY: gene
; LOCATION: (2191)..(6098)
; OTHER INFORMATION: Polr2a
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (2291)..(2477)
; OTHER INFORMATION: Exon 26 of polr2a gene
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (2817)..(2966)
; OTHER INFORMATION: Exon 27 of polr2a gene
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (3050)..(4206)
; OTHER INFORMATION: Exon 28 of polr2a gene
; FEATURE:
; NAME/KEY: 3'UTR
; LOCATION: (4207)..(6098)
; OTHER INFORMATION: gene=Polr2a
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4469)..(5056)
; OTHER INFORMATION: Region: Internal Ribosome Entry Site (IRES)
; FEATURE:
; NAME/KEY: gene

```
; LOCATION: (5057)..(5860)
; OTHER INFORMATION: gene=neo
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (12140)..(12607)
; OTHER INFORMATION: Region: pUC vector
US-10-867-628-1

Query Match          71.2%; Score 17.8; DB 8; Length 12607;
Best Local Similarity 90.5%; Pred. No. 2.4e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      5 AAGAGCGCAGACTGGTCTTCC 25
Db      11670 AAGAGCGCAGACTGGTCTTCC 11690

RESULT 48
US-10-359-050-13
; Sequence 13, Application US/10359050
; Publication No. US20030186291A1
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS DEUTSCHLAND GMBH
; TITLE OF INVENTION: GENETICALLY ENGINEERED PHIC31-INTERGASE GENES
; FILE REFERENCE: AR03-001
; CURRENT APPLICATION NUMBER: US/10/359,050
; CURRENT FILING DATE: 2003-02-05
; PRIOR APPLICATION NUMBER: US 60/354,741
; PRIOR FILING DATE: 2002-02-06
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 13
; LENGTH: 12645
; TYPE: DNA
; ORGANISM: pROSA-SA-C31-Int (CNLS) -CO
US-10-359-050-13

Query Match          71.2%; Score 17.8; DB 6; Length 12645;
Best Local Similarity 90.5%; Pred. No. 2.4e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      5 AAGAGCGCAGACTGGTCTTCC 25
Db      7108 AAGAGCGCAGACTGGTCTTCC 7128

RESULT 49
US-10-685-837-1
; Sequence 1, Application US/10685837
; Publication No. US20050071893A1
; GENERAL INFORMATION:
; APPLICANT: Seibler, Jost
; APPLICANT: 'Schwenk, Frieder
; APPLICANT: 'Kuhn, Ralf
; APPLICANT: Kuter-Luks, Birgit
; TITLE OF INVENTION: siRNA mediated gene silencing in transgenic animals
; FILE REFERENCE: 022698us JH/BM
; CURRENT APPLICATION NUMBER: US/10/685,837
; CURRENT FILING DATE: 2003-10-15
; PRIOR APPLICATION NUMBER: US60/420,476
; PRIOR FILING DATE: 2002-10-22
; PRIOR APPLICATION NUMBER: US60/467,814
; PRIOR FILING DATE: 2003-05-03
; PRIOR APPLICATION NUMBER: US60/485,969
; PRIOR FILING DATE: 2003-07-03
; PRIOR APPLICATION NUMBER: 02023283.1
; NUMBER OF SEQ ID NOS: 220
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 13139
; TYPE: DNA
; ORGANISM: Artificial Sequence

; LOCATION: (5057)..(5860)
; OTHER INFORMATION: Description of Artificial Sequence: Rosa26 locus
; OTHER INFORMATION: sequence
US-10-685-837-1

Query Match          71.2%; Score 17.8; DB 9; Length 13139;
Best Local Similarity 90.5%; Pred. No. 2.4e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      5 AAGAGCGCAGACTGGTCTTCC 25
Db      4940 AAGAGCACAGACTGCTCTTCC 4960

RESULT 50
US-10-685-837-4
; Sequence 4, Application US/10685837
; Publication No. US20050071893A1
; GENERAL INFORMATION:
; APPLICANT: Seibler, Jost
; APPLICANT: 'Schwenk, Frieder
; APPLICANT: Kuhn, Ralf
; APPLICANT: Kuter-Luks, Birgit
; TITLE OF INVENTION: siRNA mediated gene silencing in transgenic animals
; FILE REFERENCE: 022698us JH/BM
; CURRENT APPLICATION NUMBER: US/10/685,837
; CURRENT FILING DATE: 2003-10-15
; PRIOR APPLICATION NUMBER: US60/420,476
; PRIOR FILING DATE: 2002-10-22
; PRIOR APPLICATION NUMBER: US60/467,814
; PRIOR FILING DATE: 2003-05-03
; PRIOR APPLICATION NUMBER: US60/485,969
; PRIOR FILING DATE: 2003-07-03
; PRIOR APPLICATION NUMBER: 02023283.1
; NUMBER OF SEQ ID NOS: 220
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 14947
; TYPE: DNA
; ORGANISM: Artificial Sequence

; LOCATION: (5057)..(5860)
; OTHER INFORMATION: Description of Artificial Sequence: Targeting
; OTHER INFORMATION: vector for Rosa26 locus with a Fluc-hygro insert
US-10-685-837-4

Query Match          71.2%; Score 17.8; DB 9; Length 14947;
Best Local Similarity 90.5%; Pred. No. 2.4e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      5 AAGAGCGCAGACTGGTCTTCC 25
Db      6824 AAGAGCACAGACTGCTCTTCC 6844

Search completed: February 3, 2006, 15:44:12
Job time : 372.556 secs
```


GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 3, 2006, 11:35:31 ; Search time 343.111 Seconds
(without alignments)
61.064 Million cell updates/sec

Title: US-10-719-900-35

Perfect score: 25

Sequence: 1 aaaaagagcgagactgtcttc 25

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 6068529 seqs, 419036697 residues

Total number of hits satisfying chosen parameters: 12137058

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database : Published Applications NA New.*

- 1: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
- 2: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
- 5: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
- 7: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
- 9: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
- 10: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
- 11: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	18.8	75.2	43445	8	US-11-124-020A-1
2	18.8	75.2	43445	8	US-11-124-020A-2
3	18.2	72.8	3161	7	US-10-750-185-32958
4	18.2	72.8	3161	7	US-10-750-623-32958
5	18.2	72.8	7715	8	US-11-147-109-1
6	17.8	71.2	95	7	US-10-909-125-1605
7	17.8	71.2	2576	8	US-11-136-527-3672
8	17.8	71.2	2712	8	US-11-136-527-2401
9	17.8	71.2	171936	7	US-10-933-025-24
10	17.6	70.4	580	8	US-11-128-061-1155
11	17.6	70.4	580	8	US-11-128-061-4797
12	17.6	70.4	580	8	US-11-128-049-1155
13	17.6	70.4	580	8	US-11-128-049-4797
14	17.6	70.4	2180	7	US-10-750-185-48791
15	17.6	70.4	2180	7	US-10-750-623-48791
16	17.6	70.4	2298	8	US-11-037-243-21
17	17.2	68.8	1574	7	US-10-750-185-33387
18	17.2	68.8	1574	7	US-10-750-623-33387
19	17	68.0	201	7	US-10-995-561-76336
20	17	68.0	1630	8	US-11-136-527-711
21	17	68.0	2576	7	US-10-750-185-62716
22	17	68.0	2576	7	US-10-750-623-62716

Sequence 43029, A
Sequence 43029, A
Sequence 13473, A
Sequence 1233, Ap
Sequence 45468, A
Sequence 45468, A
Sequence 1621, Ap
Sequence 1621, Ap
Sequence 5263, Ap
Sequence 5263, Ap
Sequence 51, Appl
Sequence 51, Appl
Sequence 51696, A
Sequence 51696, A
Sequence 21824, A
Sequence 25, Appl
Sequence 32523, A
Sequence 32523, A
Sequence 57056, A
Sequence 57056, A
Sequence 65863, A
Sequence 65863, A
Sequence 6964, Ap
Sequence 6964, Ap
Sequence 3322, Ap
Sequence 3322, Ap
Sequence 6964, Ap
Sequence 6964, Ap
Sequence 4682, Ap
Sequence 4682, Ap
Sequence 584, App
Sequence 1040, Ap
Sequence 1040, Ap
Sequence 2708, Ap
Sequence 2708, Ap
Sequence 6350, Ap
Sequence 6350, Ap
Sequence 2708, Ap
Sequence 6350, Ap
Sequence 1928, Ap
Sequence 1928, Ap
Sequence 5570, Ap
Sequence 5570, Ap
Sequence 6661, Ap
Sequence 6661, Ap
Sequence 7498, Ap
Sequence 7498, Ap
Sequence 7834, Ap
Sequence 7834, Ap
Sequence 7913, Ap
Sequence 7913, Ap
Sequence 8182, Ap
Sequence 8182, Ap
Sequence 710, App
Sequence 710, App
Sequence 3817, Ap
Sequence 3817, Ap
Sequence 7985, Ap
Sequence 7985, Ap
Sequence 3387, Ap
Sequence 3387, Ap
Sequence 3387, Ap
Sequence 3387, Ap
Sequence 3565, Ap
Sequence 3565, Ap
Sequence 4091, Ap
Sequence 4091, Ap
Sequence 3738, Ap
Sequence 3738, Ap
Sequence 3527, Ap
Sequence 3527, Ap
Sequence 3005, Ap
Sequence 3005, Ap
Sequence 3402, Ap
Sequence 3402, Ap
Sequence 1929, Ap
Sequence 1929, Ap
Sequence 606, App
Sequence 606, App
Sequence 4086, Ap
Sequence 4086, Ap
Sequence 53, Appl
Sequence 53, Appl
Sequence 3390, Ap
Sequence 3390, Ap
Sequence 3390, Ap
Sequence 3390, Ap
Sequence 3, Appl
Sequence 3, Appl
Sequence 3985, Ap
Sequence 3985, Ap
Sequence 3585, Ap
Sequence 3585, Ap
Sequence 4042, Ap
Sequence 4042, Ap
Sequence 3889, Ap
Sequence 3889, Ap
Sequence 728, App
Sequence 728, App
Sequence 13379, A
Sequence 13379, A
Sequence 3897, Ap
Sequence 3897, Ap
Sequence 3897, Ap
Sequence 3897, Ap
Sequence 4054, Ap
Sequence 4054, Ap

96	16.2	64.8	17517	8	US-11-136-527-3650	Sequence 3650, Ap
97	16.2	64.8	17517	8	US-11-136-527-3650	Sequence 3650, Ap
98	16.2	64.8	98716	7	US-10-995-561-13531	Sequence 13331, A
99	16.2	64.8	144524	7	US-11-004-762-16	Sequence 16, Appl
100	16.2	64.8	171936	7	US-10-933-025-24	Sequence 24, Appl
101	16	64.0	384	8	US-11-043-788-13	Sequence 13, Appl
102	16	64.0	705	6	US-10-927-322-2	Sequence 2, Appl
103	16	64.0	922	7	US-10-750-185-37566	Sequence 37566, A
104	16	64.0	922	7	US-10-750-623-37566	Sequence 37566, A
105	16	64.0	995	7	US-10-750-185-49022	Sequence 49022, A
106	16	64.0	995	7	US-10-750-623-49022	Sequence 49022, A
107	16	64.0	1061	8	US-11-043-788-12	Sequence 12, Appl
108	16	64.0	1078	7	US-10-750-185-51934	Sequence 51934, A
109	16	64.0	1078	7	US-10-750-623-51934	Sequence 51934, A
110	16	64.0	1113	7	US-10-750-185-55515	Sequence 55515, A
111	16	64.0	1113	7	US-10-750-623-55515	Sequence 55515, A
112	16	64.0	1150	8	US-11-043-788-11	Sequence 11, Appl
113	16	64.0	1296	7	US-10-467-657-7233	Sequence 7233, Ap
114	16	64.0	1341	8	US-11-136-527-2929	Sequence 2929, Ap
115	16	64.0	1341	8	US-11-136-527-7025	Sequence 7025, Ap
116	16	64.0	1566	7	US-10-467-657-7237	Sequence 7237, Ap
117	16	64.0	1742	7	US-10-750-185-36588	Sequence 36588, A
118	16	64.0	1742	7	US-10-750-623-36588	Sequence 36588, A
119	16	64.0	2593	7	US-10-750-185-58794	Sequence 58794, A
120	16	64.0	2593	7	US-10-750-623-58794	Sequence 58794, A
121	16	64.0	3323	7	US-10-750-185-60084	Sequence 60084, A
122	16	64.0	3323	7	US-10-750-623-60084	Sequence 60084, A
123	16	64.0	4016	8	US-11-043-788-10	Sequence 10, Appl
124	16	64.0	10129	8	US-11-044-111-21	Sequence 21, Appl
125	16	64.0	92584	8	US-11-117-187-194	Sequence 194, Appl
126	16	64.0	137935	7	US-10-995-561-13278	Sequence 13278, A
127	16	64.0	173602	8	US-11-121-086-25	Sequence 25, Appl
128	16	64.0	197781	8	US-11-112-908-34	Sequence 34, Appl
129	16	64.0	198161	7	US-10-775-169-52	Sequence 52, Appl
130	16	64.0	198285	7	US-10-775-169-338	Sequence 338, Appl
131	16	64.0	217623	8	US-11-112-908-33	Sequence 33, Appl
132	16	64.0	1691140	8	US-11-091-018-1	Sequence 1, Appl
133	15.8	63.2	201	7	US-10-995-561-32198	Sequence 32198, A
134	15.8	63.2	600	7	US-10-750-185-21237	Sequence 21237, A
135	15.8	63.2	600	7	US-10-750-623-21237	Sequence 21237, A
136	15.8	63.2	1109	7	US-10-750-185-35454	Sequence 35454, A
137	15.8	63.2	1109	7	US-10-750-623-35454	Sequence 35454, A
138	15.8	63.2	2196	7	US-10-750-185-44622	Sequence 44622, A
139	15.8	63.2	2196	7	US-10-750-623-44622	Sequence 44622, A
140	15.8	63.2	2452	8	US-11-136-527-464	Sequence 464, Appl
141	15.8	63.2	131855	8	US-11-112-908-29	Sequence 29, Appl
142	15.8	63.2	150173	8	US-11-112-908-26	Sequence 26, Appl
143	15.8	63.2	166020	8	US-11-112-908-28	Sequence 28, Appl
144	15.8	63.2	171247	8	US-11-112-908-27	Sequence 27, Appl
145	15.8	63.2	187745	8	US-11-121-086-83	Sequence 83, Appl
146	15.8	63.2	387780	7	US-10-995-561-13259	Sequence 13259, A
147	15.6	62.4	25	8	US-11-121-849-490854	Sequence 490854, A
148	15.6	62.4	201	7	US-10-995-561-44151	Sequence 44151, A
149	15.6	62.4	201	7	US-10-995-561-44375	Sequence 44375, A
150	15.6	62.4	583	8	US-11-128-061-2491	Sequence 2491, Ap

ALIGNMENTS

RESULT 1

US-11-124-020A-1

Sequence 1, Application US/11124020A

Publication No. US20050287558A1

GENERAL INFORMATION:

APPLICANT: Rosanne M. Crooke

APPLICANT: Steven Mah

APPLICANT: Mark J. Graham

TITLE OF INVENTION: SNPS OF APOLIPOPROTEIN B AND MODULATION

TITLE OF INVENTION: OF THEIR EXPRESSION

FILE REFERENCE: BIO0021US

CURRENT APPLICATION NUMBER: US/11/124,020A

CURRENT FILING DATE: 2005-05-05

Query Match 75.2%; Score 18.8; DB 8; Length 43445;

Best Local Similarity 90.9%; Pred. No. 25;

Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAAGAGCGCAGACTGGTCT 22

Db 17678 AAAAAAGAGCGTACTGGTCT 17657

RESULT 2

US-11-124-020A-2/c

Sequence 2, Application US/11124020A

Publication No. US20050287558A1

GENERAL INFORMATION:

APPLICANT: Rosanne M. Crooke

APPLICANT: Mark J. Graham

APPLICANT: Steven Mah

TITLE OF INVENTION: SNPS OF APOLIPOPROTEIN B AND MODULATION

TITLE OF INVENTION: OF THEIR EXPRESSION

FILE REFERENCE: BIO0021US

CURRENT APPLICATION NUMBER: US/11/124,020A

CURRENT FILING DATE: 2005-05-05

Query Match 75.2%; Score 18.8; DB 8; Length 43445;

Best Local Similarity 90.9%; Pred. No. 25;

Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAAAAGAGCGCAGACTGGTCT 22

Db 17678 AAAAAAGAGCGTACTGGTCT 17657

RESULT 3

US-10-750-185-32958/c

Sequence 32958, Application US/10750185

Publication No. US20050260603A1

GENERAL INFORMATION:

APPLICANT: MMI GENOMICS, INC.

APPLICANT: Denise, Sue K.

APPLICANT: KERR, Richard

APPLICANT: ROSENFELD, David

APPLICANT: HOLM, Tom

APPLICANT: BATES, Stephen

APPLICANT: FANTIN, Dennis

TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS

FILE REFERENCE: MM1100-2

CURRENT APPLICATION NUMBER: US/10/750,185

CURRENT FILING DATE: 2003-12-31

PRIOR APPLICATION NUMBER: US 60/437,482

PRIOR FILING DATE: 2002-12-31

NUMBER OF SEQ ID NOS: 64922

SOFTWARE: PatentIN version 3.1

SEQ ID NO 32958

LENGTH: 3161


```
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1147)..(1147)
; OTHER INFORMATION: n is a, c, g, or t
US-11-136-527-3672

Query Match          71.2%; Score 17.8; DB 8; Length 2576;
Best Local Similarity 90.5%; Pred. No. 48;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 AAGAGCGCAGACTGGTCTTCC 25
    ||||| ||||| ||||| |||||
Db 814 AAGAGCACAGACTGCTCTTCC 834

RESULT 8
US-11-136-527-2401
; Sequence 2401, Application US/111136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2401
; LENGTH: 2712
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-2401

Query Match          71.2%; Score 17.8; DB 8; Length 2712;
Best Local Similarity 90.5%; Pred. No. 48;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 AAGAGCGCAGACTGGTCTTCC 25
    ||||| ||||| ||||| |||||
Db 87 AAGAGCAGTGGTCTTCC 107

RESULT 9
US-10-933-025-24
; Sequence 24, Application US/10933025
; Publication No. US20050265967A1
; GENERAL INFORMATION:
; APPLICANT: ROSEN, STEVEN
; APPLICANT: HEMMERICH, STEFAN
; APPLICANT: TOMITA, MEGUMI
; TITLE OF INVENTION: Sulfotransferases and methods of use
; FILE REFERENCE: UCAL-230CON
; CURRENT APPLICATION NUMBER: US/10/933,025
; CURRENT FILING DATE: 2004-09-01
; PRIOR APPLICATION NUMBER: 10/025,966
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/258,577
; PRIOR FILING DATE: 2000-12-27
; PRIOR APPLICATION NUMBER: 60/267,831
; PRIOR FILING DATE: 2001-09-02
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24
; LENGTH: 171936
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-933-025-24

; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1147)..(1147)
; OTHER INFORMATION: n is a, c, g, or t
US-11-136-527-3672

Query Match          71.2%; Score 17.8; DB 7; Length 171936;
Best Local Similarity 90.5%; Pred. No. 91;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 AAGAGCGCAGACTGGTCTTCC 25
    ||||| ||||| ||||| |||||
Db 144637 AAGAGCACAGACTGCTCTTCC 144657

RESULT 10
US-11-128-061-1155/c
; Sequence 1155, Application US/11128061
; Publication No. US20060003958A1
; GENERAL INFORMATION:
; APPLICANT: Melville, Mark W.
; APPLICANT: Charlebois, Timothy S.
; APPLICANT: Mounts, William M.
; APPLICANT: Hann, Louane E.
; APPLICANT: Sinacore, Martin S.
; APPLICANT: Leonard, Mark W.
; APPLICANT: Brown, Eugene L.
; APPLICANT: Miller, Christopher P.
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES RELATED TO OLIGONUCLEOTIDE ARRAYS
; FILE REFERENCE: 01997.027701
; CURRENT APPLICATION NUMBER: US/11/128,061
; CURRENT FILING DATE: 2005-05-11
; PRIOR APPLICATION NUMBER: US 60/570,425
; PRIOR FILING DATE: 2004-05-11
; NUMBER OF SEQ ID NOS: 7285
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1155
; LENGTH: 580
; TYPE: DNA
; ORGANISM: Cricetulus griseus
US-11-128-061-1155

Query Match          70.4%; Score 17.6; DB 8; Length 580;
Best Local Similarity 83.3%; Pred. No. 47;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAGAGCGCAGACTGGTCTTC 24
    ||||| ||||| ||||| |||||
Db 53 AAAAAGAGCACAGACTAGTCTTC 30

RESULT 11
US-11-128-061-4797/c
; Sequence 4797, Application US/11128061
; Publication No. US20060003958A1
; GENERAL INFORMATION:
; APPLICANT: Melville, Mark W.
; APPLICANT: Charlebois, Timothy S.
; APPLICANT: Mounts, William M.
; APPLICANT: Hann, Louane E.
; APPLICANT: Sinacore, Martin S.
; APPLICANT: Leonard, Mark W.
; APPLICANT: Brown, Eugene L.
; APPLICANT: Miller, Christopher P.
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES RELATED TO OLIGONUCLEOTIDE ARRAYS
; FILE REFERENCE: 01997.027701
; CURRENT APPLICATION NUMBER: US/11/128,061
; CURRENT FILING DATE: 2005-05-11
; PRIOR APPLICATION NUMBER: US 60/570,425
; PRIOR FILING DATE: 2004-05-11
; NUMBER OF SEQ ID NOS: 7285
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 4797
; LENGTH: 580
; TYPE: DNA
; ORGANISM: Cricetulus griseus
```


US-11-128-061-4797

Query Match 70.4%; Score 17.6; DB 8; Length 580;
Best Local Similarity 83.3%; Pred. No. 47;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAAGCGCAGACTGGTCTTC 24
||||| ||| ||||| ||||| |||||
DB 53 AAAAAACAGCAGACTAGTCTTC 30

RESULT 12

US-11-128-049-1155/c
; Sequence 1155, Application US/11128049
; Publication No. US20060010513A1
; GENERAL INFORMATION:
; APPLICANT: Melville, Mark W.
; APPLICANT: Charlebois, Timothy S.
; APPLICANT: Mounts, William M.
; APPLICANT: Hann, Louane E.
; APPLICANT: Sinacore, Martin S.
; APPLICANT: Leonard, Mark W.
; APPLICANT: Brown, Eugene L.
; APPLICANT: Miller, Christopher P.
; TITLE OF INVENTION: OLIGONUCLEOTIDE ARRAYS TO MONITOR GENE EXPRESSION AND METHODS FOR
; TITLE OF INVENTION: MAKING AND USING SAME
; FILE REFERENCE: 01997.027700
; CURRENT APPLICATION NUMBER: US/11/128,049
; CURRENT FILING DATE: 2005-05-11
; PRIOR APPLICATION NUMBER: US 60/570,425
; PRIOR FILING DATE: 2004-05-11
; NUMBER OF SEQ ID NOS: 7285
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1155
; LENGTH: 580
; TYPE: DNA
; ORGANISM: Cricetulus griseus
US-11-128-049-1155

Query Match 70.4%; Score 17.6; DB 8; Length 580;
Best Local Similarity 83.3%; Pred. No. 47;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAAGCGCAGACTGGTCTTC 24
||||| ||| ||||| ||||| |||||
DB 53 AAAAAACAGCAGACTAGTCTTC 30

RESULT 13

US-11-128-049-4797/c
; Sequence 4797, Application US/11128049
; Publication No. US20060010513A1
; GENERAL INFORMATION:
; APPLICANT: Melville, Mark W.
; APPLICANT: Charlebois, Timothy S.
; APPLICANT: Mounts, William M.
; APPLICANT: Hann, Louane E.
; APPLICANT: Sinacore, Martin S.
; APPLICANT: Leonard, Mark W.
; APPLICANT: Brown, Eugene L.
; APPLICANT: Miller, Christopher P.
; TITLE OF INVENTION: OLIGONUCLEOTIDE ARRAYS TO MONITOR GENE EXPRESSION AND METHODS FOR
; TITLE OF INVENTION: MAKING AND USING SAME
; FILE REFERENCE: 01997.027700
; CURRENT APPLICATION NUMBER: US/11/128,049
; CURRENT FILING DATE: 2005-05-11
; PRIOR APPLICATION NUMBER: US 60/570,425
; PRIOR FILING DATE: 2004-05-11
; NUMBER OF SEQ ID NOS: 7285
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 4797
; LENGTH: 580
; TYPE: DNA

; ORGANISM: Cricetulus griseus
US-11-128-049-4797

Query Match 70.4%; Score 17.6; DB 8; Length 580;
Best Local Similarity 83.3%; Pred. No. 47;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAAGCGCAGACTGGTCTTC 24
||||| ||| ||||| ||||| |||||
DB 53 AAAAAACAGCAGACTAGTCTTC 30

RESULT 14

US-10-750-185-48791/c
; Sequence 48791, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 48791
; LENGTH: 2180
; TYPE: DNA
; ORGANISM: Bovine 19866881276311
US-10-750-185-48791

Query Match 70.4%; Score 17.6; DB 7; Length 2180;
Best Local Similarity 83.3%; Pred. No. 58;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAAGCGCAGACTGGTCTTC 24
||||| ||| ||||| ||||| |||||
DB 1120 AGAACAGAGTACAGACTGGTCTTC 1097

RESULT 15

US-10-750-623-48791/c
; Sequence 48791, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 48791
; LENGTH: 2180
; TYPE: DNA
; ORGANISM: Bovine 19866881276311
US-10-750-623-48791

Query Match 70.4%; Score 17.6; DB 7; Length 2180;
 Best Local Similarity 83.3%; Pred. No. 58;
 Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAAGAGCGCAGACTGGTCTTC 24
 DB 1120 AGAACAGAGTACAGACTGGTCTTC 1097

RESULT 16
 US-11-037-243-21/c
 ; Sequence 21, Application US/11037243
 ; Publication No. US20050287546A1
 ; GENERAL INFORMATION:
 ; APPLICANT: FLOWMAN, GREGORY
 ; APPLICANT: WHITE, DAVID
 ; APPLICANT: CAENEPEEL, SEAN
 ; APPLICANT: CHARYDCZAK, GLEN
 ; APPLICANT: MANNING, GERARD
 ; APPLICANT: SUDARSANAM, SUCHA
 ; TITLE OF INVENTION: NOVEL PROTEASES
 ; FILE REFERENCE: 038602/1214
 ; CURRENT APPLICATION NUMBER: US/11/037,243
 ; PRIOR FILING DATE: 2005-05-26
 ; PRIOR APPLICATION NUMBER: US/09/888,615
 ; PRIOR FILING DATE: 2001-06-26
 ; PRIOR APPLICATION NUMBER: 60/214,047
 ; PRIOR FILING DATE: 2000-06-26
 ; NUMBER OF SEQ ID NOS: 150
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 21
 ; LENGTH: 2298
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-11-037-243-21

Query Match 70.4%; Score 17.6; DB 8; Length 2298;
 Best Local Similarity 83.3%; Pred. No. 59;
 Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AAAAAAGAGCGCAGACTGGTCTTC 25
 DB 773 AGAACAGAGCCAGACTGGTCCACC 750

RESULT 17
 US-10-750-185-33387
 ; Sequence 33387, Application US/10750185
 ; Publication No. US20050260603A1
 ; GENERAL INFORMATION:
 ; APPLICANT: MMI GENOMICS, INC.
 ; APPLICANT: DENISE, Sue K.
 ; APPLICANT: KERR, Richard
 ; APPLICANT: ROSENFELD, David
 ; APPLICANT: HOLM, Tom
 ; APPLICANT: BATES, Stephen
 ; APPLICANT: FANTIN, Dennis
 ; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
 ; FILE REFERENCE: MM1100-2
 ; CURRENT APPLICATION NUMBER: US/10/750,185
 ; CURRENT FILING DATE: 2003-12-31
 ; PRIOR APPLICATION NUMBER: US 60/437,482
 ; PRIOR FILING DATE: 2002-12-31
 ; NUMBER OF SEQ ID NOS: 64922
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 33387
 ; LENGTH: 1574
 ; TYPE: DNA
 ; ORGANISM: Bovine 19866881632003
 US-10-750-185-33387

Query Match 68.8%; Score 17.2; DB 7; Length 1574;
 Best Local Similarity 86.4%; Pred. No. 85;

Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 3 AAAAGAGCGCAGACTGGTCTTC 24
 DB 499 AAATGAGCCACACTGGTCTTC 520

RESULT 18
 US-10-750-623-33387
 ; Sequence 33387, Application US/10750623
 ; Publication No. US20050287531A1
 ; GENERAL INFORMATION:
 ; APPLICANT: MMI GENOMICS, INC.
 ; APPLICANT: DENISE, Sue K.
 ; APPLICANT: KERR, Richard
 ; APPLICANT: ROSENFELD, David
 ; APPLICANT: HOLM, Tom
 ; APPLICANT: BATES, Stephen
 ; APPLICANT: FANTIN, Dennis
 ; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
 ; FILE REFERENCE: MM1100-1
 ; CURRENT APPLICATION NUMBER: US/10/750,623
 ; CURRENT FILING DATE: 2003-12-31
 ; PRIOR APPLICATION NUMBER: US 60/437,482
 ; PRIOR FILING DATE: 2002-12-31
 ; NUMBER OF SEQ ID NOS: 64922
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 33387
 ; LENGTH: 1574
 ; TYPE: DNA
 ; ORGANISM: Bovine 19866881632003
 US-10-750-623-33387

Query Match 68.8%; Score 17.2; DB 7; Length 1574;
 Best Local Similarity 86.4%; Pred. No. 85;
 Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 AAAAGAGCGCAGACTGGTCTTC 24
 DB 499 AAATGAGCCACACTGGTCTTC 520

RESULT 19
 US-10-995-561-76336/c
 ; Sequence 76336, Application US/10995561
 ; Publication No. US2005027054A1
 ; GENERAL INFORMATION:
 ; APPLICANT: CARGILL, Michele et al.
 ; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
 ; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
 ; TITLE OF INVENTION: DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001559
 ; CURRENT APPLICATION NUMBER: US/10/995,561
 ; CURRENT FILING DATE: 2004-11-24
 ; NUMBER OF SEQ ID NOS: 85702
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 76336
 ; LENGTH: 201
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-995-561-76336

Query Match 68.0%; Score 17; DB 7; Length 201;
 Best Local Similarity 80.0%; Pred. No. 77;
 Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAAAGAGCGCAGACTGGTCTTC 25
 DB 30 AACAAAGAGGCGAGCTCCACTTCC 6

RESULT 20
 US-11-136-527-711/c

```
; Sequence 711, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 711
; LENGTH: 1630
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-711

Query Match      68.0%; Score 17; DB 8; Length 1630;
Best Local Similarity 80.0%; Pred. No. 1.1e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAAAGAGCGCAGACTGGTCTTCC 25
   ||||| ||||| ||||| ||||| |||||
Db 159 AAGAAATAGCTCAACGGGCTTCC 135

RESULT 21
US-10-750-185-62716
; Sequence 62716, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 62716
; LENGTH: 2576
; TYPE: DNA
; ORGANISM: Bovine 19866880397212
US-10-750-185-62716

Query Match      68.0%; Score 17; DB 7; Length 2576;
Best Local Similarity 80.0%; Pred. No. 1.1e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAAAGAGCGCAGACTGGTCTTCC 25
   ||||| ||||| ||||| ||||| |||||
Db 2387 AAAAAAGTGACGATACGTCCTCC 2411

RESULT 22
US-10-750-623-62716
; Sequence 62716, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 62716
; LENGTH: 2576
; TYPE: DNA
; ORGANISM: Bovine 19866880397212
US-10-750-623-62716

Query Match      68.0%; Score 17; DB 7; Length 2576;
Best Local Similarity 80.0%; Pred. No. 1.1e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAAAGAGCGCAGACTGGTCTTCC 25
   ||||| ||||| ||||| ||||| |||||
Db 652 AAAAAAGCAAGAGTGATCCTCC 628

RESULT 24
US-10-750-623-43029/c
; Sequence 43029, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 43029
; LENGTH: 2929
; TYPE: DNA
; ORGANISM: Bovine 19866880892057
US-10-750-185-43029

Query Match      68.0%; Score 17; DB 7; Length 2929;
Best Local Similarity 80.0%; Pred. No. 1.2e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAAAGAGCGCAGACTGGTCTTCC 25
   ||||| ||||| ||||| ||||| |||||
Db 652 AAAAAAGCAAGAGTGATCCTCC 628

RESULT 24
US-10-750-623-43029/c
; Sequence 43029, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 62716
; LENGTH: 2576
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-711

Query Match      68.0%; Score 17; DB 8; Length 1630;
Best Local Similarity 80.0%; Pred. No. 1.1e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAAAGAGCGCAGACTGGTCTTCC 25
   ||||| ||||| ||||| ||||| |||||
Db 159 AAGAAATAGCTCAACGGGCTTCC 135

RESULT 21
US-10-750-185-62716
; Sequence 62716, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 62716
; LENGTH: 2576
; TYPE: DNA
; ORGANISM: Bovine 19866880397212
US-10-750-185-62716

Query Match      68.0%; Score 17; DB 7; Length 2576;
Best Local Similarity 80.0%; Pred. No. 1.1e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAAAGAGCGCAGACTGGTCTTCC 25
   ||||| ||||| ||||| ||||| |||||
Db 2387 AAAAAAGTGACGATACGTCCTCC 2411

RESULT 22
US-10-750-623-62716
; Sequence 62716, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 62716
; LENGTH: 2576
; TYPE: DNA
; ORGANISM: Bovine 19866880397212
US-10-750-623-62716

Query Match      68.0%; Score 17; DB 7; Length 2576;
Best Local Similarity 80.0%; Pred. No. 1.1e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAAAGAGCGCAGACTGGTCTTCC 25
   ||||| ||||| ||||| ||||| |||||
Db 652 AAAAAAGCAAGAGTGATCCTCC 628

RESULT 24
US-10-750-623-43029/c
; Sequence 43029, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 43029
; LENGTH: 2929
; TYPE: DNA
; ORGANISM: Bovine 19866880892057
US-10-750-185-43029

Query Match      68.0%; Score 17; DB 7; Length 2929;
Best Local Similarity 80.0%; Pred. No. 1.2e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAAAGAGCGCAGACTGGTCTTCC 25
   ||||| ||||| ||||| ||||| |||||
Db 652 AAAAAAGCAAGAGTGATCCTCC 628

RESULT 24
US-10-750-623-43029/c
; Sequence 43029, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 62716
; LENGTH: 2576
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-711
```

```
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 43029
; LENGTH: 2929
; TYPE: DNA
; ORGANISM: Bovine 19866880892057
US-10-750-623-43029

Query Match      68.0%; Score 17; DB 7; Length 2929;
Best Local Similarity 80.0%; Pred. No. 1.2e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAAAGAGCGCAGACTGGTCTTCC 25
   ||||| ||||| ||||| ||||| |||||
Db 652 AAAAAAGAGCCAGAGTGATCTTCC 628

RESULT 25
US-10-995-561-13473/c
; Sequence 13473, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13473
; LENGTH: 394468
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..-(394468)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-995-561-13473

Query Match      68.0%; Score 17; DB 7; Length 394468;
Best Local Similarity 80.0%; Pred. No. 2.3e+02;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAAAAAGAGCGCAGACTGGTCTTCC 25
   ||||| ||||| ||||| ||||| |||||
Db 327726 AACAAAGAGCGCAGACTCCACTTCC 327702

RESULT 26
US-11-000-688-1233
; Sequence 1233, Application US/11000688
; Publication No. US20050287544A1
; GENERAL INFORMATION:
; APPLICANT: BERTUCCI, Francois
; APPLICANT: HOULGATTE, Remi
; APPLICANT: BIRNBAUM, Daniel
; TITLE OF INVENTION: GENE EXPRESSION PROFILING OF COLON CANCER WITH DNA ARRAYS
; FILE REFERENCE: 1423-R-03
; CURRENT APPLICATION NUMBER: US/11/000,688
; CURRENT FILING DATE: 2004-12-01
; PRIOR APPLICATION NUMBER: US 60/525,987
; PRIOR FILING DATE: 2003-12-01
; NUMBER OF SEQ ID NOS: 1596
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1233
; LENGTH: 481

; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 43029
; LENGTH: 2929
; TYPE: DNA
; ORGANISM: Bovine 19866880892057
US-10-750-623-43029

Query Match      67.2%; Score 16.8; DB 8; Length 481;
Best Local Similarity 90.0%; Pred. No. 1.1e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AAAAAAGCGCAGACTGGTCTT 21
   ||||| ||||| ||||| ||||| |||||
Db 85 AAAAAAGCAGACAGACTGGTC 104

RESULT 27
US-10-185-45468/c
; Sequence 45468, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 45468
; LENGTH: 1897
; TYPE: DNA
; ORGANISM: Bovine 19866880485767
US-10-185-45468

Query Match      67.2%; Score 16.8; DB 7; Length 1897;
Best Local Similarity 90.0%; Pred. No. 1.4e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 AAGAGCGCAGACTGGTCTT 23
   ||||| ||||| ||||| ||||| |||||
Db 900 AAAAAAGTCAGACTGGTCTT 881

RESULT 28
US-10-750-623-45468/c
; Sequence 45468, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
```

```

; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 45468
; LENGTH: 1897
; TYPE: DNA
; ORGANISM: Bovine
US-10-750-623-45468
19866880485767

```

Query Match 67.2%; Score 16.8; DB 7; Length 1897;
Best Local Similarity 90.0%; Pred. No. 1.4e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0

Qy 4 AAAGAGCGCAGACTGGTCTT 23
||| ||| ||| ||| ||| ||| ||| |||
Db 900 AAAAAGTGCAGACTGGTCTT 881

RESULT 29

```

RESUMI 23
US-11-128-061-1621/c
; Sequence 1621, Application US/11128061
; Publication No. US20060003958A1
; GENERAL INFORMATION:
; APPLICANT: Melville, Mark W.
; APPLICANT: Charlebois, Timothy S.
; APPLICANT: Mounts, William M.
; APPLICANT: Hann, Louane E.
; APPLICANT: Sinacore, Martin S.
; APPLICANT: Leonard, Mark W.
; APPLICANT: Brown, Eugene L.
; APPLICANT: Miller, Christopher P.
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES RELATED TO OLIGONUCLEOTIDE ARRAYS
; TITLE OF INVENTION: TO MONITOR GENE EXPRESSION
; FILE REFERENCE: 01997.027701
; CURRENT APPLICATION NUMBER: US/11/128,061
; CURRENT FILING DATE: 2005-05-11
; PRIOR APPLICATION NUMBER: US 60/570,425
; PRIOR FILING DATE: 2004-05-11
; NUMBER OF SEQ ID NOS: 7285
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1621
; LENGTH: 578
; TYPE: DNA
; ORGANISM: Cricetulus griseus
US-11-128-061-1621

```

Query Match	66.4%;	Score 16.6;	DB 8;	Length 578;
Best Local Similarity	82.6%;	Pred. No. 1.4e+02;		
Matches 19:	Conservative	0;	Mismatches 4;	Indels 0;
	Gaps	0;		

QY 3 AAAAGAGCGCAGACTGGTCTTCC 25
||| | | | | | | | | |
Db 66 AAGAGGGTGCAGATTGGTCTTCC 44

RESULT 30

```

US-11-128-061-5263/c
; Sequence 5263, Application US/11128061
; Publication NO. US20060003958A1
; GENERAL INFORMATION:
; APPLICANT: Melville, Mark W.
; APPLICANT: Charlesbois, Timothy S.
; APPLICANT: Mounts, William M.
; APPLICANT: Hann, Louane E.
; APPLICANT: Sinacore, Martin S.
; APPLICANT: Leonard, Mark W.
; APPLICANT: Brown, Eugene L.
; APPLICANT: Miller, Christopher P.
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES RELATED TO OLIGONUCLEOTIDE ARRAYS
; TITLE OF INVENTION: TO MONITOR GENE EXPRESSION
; FILE REFERENCE: 01997.027701
; CURRENT APPLICATION NUMBER: US/11/128,061
; CURRENT FILING DATE: 2005-05-11
; PRIOR APPLICATION NUMBER: US 60/570,425

```

```

; PRIOR FILING DATE: 2004-05-11
; NUMBER OF SEQ ID NOS: 7285
; SOFTWARE: patentin version 3.3
; SEQ ID NO 5263
; LENGTH: 578
; TYPE: DNA
; ORGANISM: Cricetulus griseus
US-11-128-061-5263

```

Query Match	66.4%	Score 16.6;	DB 8;	Length 578;
Best Local Similarity	82.6%	Pred. No. 1.4e+02;		
Matches 19;	Conservative	0;	Mismatches 4;	Indels 0;
				Gaps 0;

Qy 3 AAAAGAGCGCAGACTGGTCTTCC 25
|||
Db 66 AAGAGGGTGCAGATTGGTCTTCC 44

RESULT 31

```

RESUMI 31
US-11-128-049-1621/c
; Sequence 1621, Application US/11128049
; Publication No. US20060010513A1
; GENERAL INFORMATION:
; APPLICANT: Melville, Mark W.
; APPLICANT: Charlebois, Timothy S.
; APPLICANT: Mounts, William M.
; APPLICANT: Hann, Louane E.
; APPLICANT: Sinacore, Martin S.
; APPLICANT: Leonard, Mark W.
; APPLICANT: Brown, Eugene L.
; APPLICANT: Miller, Christopher P.
; TITLE OF INVENTION: OLIGONUCLEOTIDE A
; TITLE OF INVENTION: MAKING AND USING
; FILE REFERENCE: 01997.027700
; CURRENT APPLICATION NUMBER: US/11/128049
; CURRENT FILING DATE: 2005-05-11
; PRIOR APPLICATION NUMBER: US 60/570,414
; PRIOR FILING DATE: 2004-05-11
; NUMBER OF SEQ ID NOS: 7285
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1621
; LENGTH: 578
; TYPE: DNA
; ORGANISM: Cricetus griseus
US-11-128-049-1621

```

Query Match	66.4%	Score 16.6;	DB 8;	Length 578;
Best Local Similarity	82.6%	Pred. No. 1.4e+02;		
Matches 19: Conservative	0;	Mismatches 4;	Indels 0;	Gaps 0;

QY 3 AAAGAGCGCAGACTGGTCTTCC 25
||| ||| ||| ||| ||| ||| ||| |||
dh 56 AAGAGGGTGCAGATTGGTCTTCC 44

RESULT 32

```

RESULT 32
US-11-128-049-5263/c
; Sequence 5263, Application US/11128049
; Publication No. US20060010513A1
; GENERAL INFORMATION:
; APPLICANT: Melville, Mark W.
; APPLICANT: Charlebois, Timothy S.
; APPLICANT: Mounts, William M.
; APPLICANT: Hann, Louane E.
; APPLICANT: Sinacore, Martin S.
; APPLICANT: Leonard, Mark W.
; APPLICANT: Brown, Eugene L.
; APPLICANT: Miller, Christopher P.
; TITLE OF INVENTION: OLIGONUCLEOTIDE A
; TITLE OF INVENTION: MAKING AND USING
; FILE REFERENCE: 01997-027700
; CURRENT APPLICATION NUMBER: US/11128
; CURRENT FILING DATE: 2005-05-11

```

```
; PRIOR APPLICATION NUMBER: US 60/570,425
; PRIOR FILING DATE: 2004-05-11
; NUMBER OF SEQ ID NOS: 7285
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 5263
; LENGTH: 578
; TYPE: DNA
; ORGANISM: Cricetulus griseus
US-11-128-049-5263

Query Match      66.4%; Score 16.6; DB 8; Length 578;
Best Local Similarity 82.6%; Pred. No. 1.4e+02;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 AAAAGAGCGCAGACTGGTCTTC 25
   ||||| ||||| ||||| |||||
Db 66 AAGAGGGTGCAGATTGGTCTTC 44

RESULT 33
US-11-112-908-51
; Sequence 51, Application US/11112908
; Publication No. US20050260659A1
; GENERAL INFORMATION:
; APPLICANT: Harris, Cole
; APPLICANT: Davis, Lisa M.
; TITLE OF INVENTION: Breast Cancer Biomarkers
; FILE REFERENCE: 04-164-US
; CURRENT APPLICATION NUMBER: US/11/112,908
; PRIOR FILING DATE: 2005-04-22
; PRIOR APPLICATION NUMBER: US 60/564,758
; PRIOR FILING DATE: 2004-04-23
; PRIOR APPLICATION NUMBER: US 60/575,978
; PRIOR FILING DATE: 2004-06-01
; PRIOR APPLICATION NUMBER: US 60/631,702
; PRIOR FILING DATE: 2004-11-30
; PRIOR APPLICATION NUMBER: US 60/633,826
; PRIOR FILING DATE: 2004-12-07
; NUMBER OF SEQ ID NOS: 511
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 51
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-112-908-51

Query Match      66.4%; Score 16.6; DB 8; Length 157224;
Best Local Similarity 82.6%; Pred. No. 3.2e+02;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAGAGCGCAGACTGGTCTT 23
   ||||| | ||||| ||||| |||||
Db 75474 AAAAAAATCGCAGACTTGCCTT 75496

RESULT 34
US-11-112-908-50
; Sequence 50, Application US/11112908
; Publication No. US20050260659A1
; GENERAL INFORMATION:
; APPLICANT: Harris, Cole
; APPLICANT: Davis, Lisa M.
; TITLE OF INVENTION: Breast
; FILE REFERENCE: 04-164-US
; CURRENT APPLICATION NUMBER: US/11/112,908
; PRIOR FILING DATE: 2005-04-22
; PRIOR APPLICATION NUMBER: US 60/564,758
; PRIOR FILING DATE: 2004-04-23
; PRIOR APPLICATION NUMBER: US 60/575,978
; PRIOR FILING DATE: 2004-06-01
; PRIOR APPLICATION NUMBER: US 60/631,702
; PRIOR FILING DATE: 2004-11-30
; PRIOR APPLICATION NUMBER: US 60/633,826
```

```
; PRIOR APPLICATION NUMBER: US 60/570,425
; PRIOR FILING DATE: 2004-05-11
; NUMBER OF SEQ ID NOS: 7285
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 50
; LENGTH: 170189
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-112-908-50

Query Match      66.4%; Score 16.6; DB 8; Length 170189;
Best Local Similarity 82.6%; Pred. No. 3.3e+02;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAAAAGAGCGCAGACTGGTCTT 23
   ||||| ||||| ||||| |||||
Db 108234 AAAAAAATCGCAGACTTGCCTT 108256

RESULT 35
US-10-750-185-51696/c
; Sequence 51696, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM11100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 51696
; LENGTH: 3684
; TYPE: DNA
; ORGANISM: Bovine
US-10-750-185-51696

Query Match      65.6%; Score 16.4; DB 7; Length 3684;
Best Local Similarity 94.4%; Pred. No. 2.3e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAAAGAGCGCAGACTG 18
   ||||| ||||| ||||| |||||
Db 2655 AAAAAAGAGCAGACTG 2638

RESULT 36
US-10-750-623-51696/c
; Sequence 51696, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM11100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
```

```
; SEQ ID NO 51696
; LENGTH: 3684
; TYPE: DNA
; ORGANISM: Bovine 1986680775622
US-10-750-623-51696

Query Match          65.6%; Score 16.4; DB 7; Length 3684;
Best Local Similarity 94.4%; Pred. No. 2.3e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAAAAGAGCGCAGACTG 18
   ||||| ||||| ||||| |||||
Db 2655 AAAAAAGAGCAGACTG 2638

RESULT 37
US-10-770-726-21824
; Sequence 21824, Application US/10770726
; Publication No. US20050266409A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Brown, Eugene
; APPLICANT: Liu, Wei
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING, PREVENTING, AND TREATING
; FILE REFERENCE: AM101079 (031896-010000)
; CURRENT APPLICATION NUMBER: US/10/770,726
; CURRENT FILING DATE: 2004-02-04
; NUMBER OF SEQ ID NOS: 48640
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 21824
; LENGTH: 21
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-770-726-21824

Query Match          64.8%; Score 16.2; DB 7; Length 21;
Best Local Similarity 85.7%; Pred. No. 1.3e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 AAAAGCGCAGACTGGTCTT 23
   ||||| ||||| ||||| |||||
Db 1 AAATGAGGAGGACTGGTCTT 21

RESULT 38
US-10-502-972-25
; Sequence 25, Application US/10502972
; Publication No. US2005025573A1
; GENERAL INFORMATION:
; APPLICANT: Chambers, Ian
; APPLICANT: Smith, Austin G.
; TITLE OF INVENTION: Pluripotency determining factors and uses thereof
; FILE REFERENCE: 2325.002000/RVE/RAS
; CURRENT APPLICATION NUMBER: US/10/502,972
; CURRENT FILING DATE: 2004-07-30
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 25
; LENGTH: 196
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-502-972-25

Query Match          64.8%; Score 16.2; DB 7; Length 196;
Best Local Similarity 85.7%; Pred. No. 1.8e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 5 AAGAGCGCAGACTGGTCTTCC 25
   ||||| ||||| ||||| |||||
Db 27 AAGAGCAGTACTGCTCTTCC 47

RESULT 39
US-10-995-561-32523/c
; Sequence 32523, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 32523
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-32523

Query Match          64.8%; Score 16.2; DB 7; Length 201;
Best Local Similarity 85.7%; Pred. No. 1.8e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 AAAGAGCGCAGACTGGTCTT 24
   ||||| ||||| ||||| |||||
Db 140 AAAGAGGAGGAGCTGGCCTT 120

RESULT 40
US-10-995-561-57056/c
; Sequence 57056, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 57056
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-57056

Query Match          64.8%; Score 16.2; DB 7; Length 201;
Best Local Similarity 85.7%; Pred. No. 1.8e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 AAAAGAGCGCAGACTGGTCTT 23
   ||||| ||||| ||||| |||||
Db 149 AAATGAGGAGGAGCTGGTCTT 129

RESULT 41
US-10-995-561-65863
; Sequence 65863, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 65863
```

```
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-65863

Query Match          64.8%; Score 16.2; DB 7; Length 201;
Best Local Similarity 85.7%; Pred. No. 1.8e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 AAAAGAGCGCAGACTGGTCTTT 23
   ||||| ||||| ||||| |||||
Db 53 AAATGAGAGGAGCTGGTCTT 73

RESULT 42
US-11-128-061-3322
; Sequence 3322, Application US/11128061
; Publication No. US20060003958A1
; GENERAL INFORMATION:
; APPLICANT: Melville, Mark W.
; APPLICANT: Charlebois, Timothy S.
; APPLICANT: Mounts, William M.
; APPLICANT: Hann, Louane E.
; APPLICANT: Sinacore, Martin S.
; APPLICANT: Leonard, Mark W.
; APPLICANT: Brown, Eugene L.
; APPLICANT: Miller, Christopher P.
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES RELATED TO OLIGONUCLEOTIDE ARRAYS
; FILE REFERENCE: 01997.027701
; CURRENT APPLICATION NUMBER: US/11/128,061
; CURRENT FILING DATE: 2005-05-11
; PRIOR APPLICATION NUMBER: US 60/570,425
; PRIOR FILING DATE: 2004-05-11
; NUMBER OF SEQ ID NOS: 7285
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 3322
; LENGTH: 275
; TYPE: DNA
; ORGANISM: Cricetulus longicaudatus
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (172)..(192)
; OTHER INFORMATION: n is a, c, g, or t
US-11-128-061-3322

Query Match          64.8%; Score 16.2; DB 8; Length 275;
Best Local Similarity 85.7%; Pred. No. 1.9e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 5 AAGAGCGCAGACTGGTCTTCC 25
   ||||| ||||| ||||| |||||
Db 198 AAGAGCACTGACTGCTCTTCC 218

RESULT 43
US-11-128-061-6964
; Sequence 6964, Application US/11128061
; Publication No. US20060003958A1
; GENERAL INFORMATION:
; APPLICANT: Melville, Mark W.
; APPLICANT: Charlebois, Timothy S.
; APPLICANT: Mounts, William M.
; APPLICANT: Hann, Louane E.
; APPLICANT: Sinacore, Martin S.
; APPLICANT: Leonard, Mark W.
; APPLICANT: Brown, Eugene L.
; APPLICANT: Miller, Christopher P.
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES RELATED TO OLIGONUCLEOTIDE ARRAYS
; FILE REFERENCE: 01997.027701
; CURRENT APPLICATION NUMBER: US/11/128,061
; CURRENT FILING DATE: 2005-05-11
```

```
; PRIOR APPLICATION NUMBER: US 60/570,425
; PRIOR FILING DATE: 2004-05-11
; NUMBER OF SEQ ID NOS: 7285
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 6964
; LENGTH: 275
; TYPE: DNA
; ORGANISM: Cricetulus longicaudatus
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (172)..(192)
; OTHER INFORMATION: n is a, c, g, or t
US-11-128-061-6964

Query Match          64.8%; Score 16.2; DB 8; Length 275;
Best Local Similarity 85.7%; Pred. No. 1.9e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 5 AAGAGCGCAGACTGGTCTTCC 25
   ||||| ||||| ||||| |||||
Db 198 AAGAGCACTGACTGCTCTTCC 218

RESULT 44
US-11-128-049-3322
; Sequence 3322, Application US/11128049
; Publication No. US20060010513A1
; GENERAL INFORMATION:
; APPLICANT: Melville, Mark W.
; APPLICANT: Charlebois, Timothy S.
; APPLICANT: Mounts, William M.
; APPLICANT: Hann, Louane E.
; APPLICANT: Sinacore, Martin S.
; APPLICANT: Leonard, Mark W.
; APPLICANT: Brown, Eugene L.
; APPLICANT: Miller, Christopher P.
; TITLE OF INVENTION: OLIGONUCLEOTIDE ARRAYS TO MONITOR GENE EXPRESSION AND METHODS FOR
; FILE REFERENCE: 01997.027700
; CURRENT APPLICATION NUMBER: US/11/128,049
; CURRENT FILING DATE: 2005-05-11
; PRIOR APPLICATION NUMBER: US 60/570,425
; PRIOR FILING DATE: 2004-05-11
; NUMBER OF SEQ ID NOS: 7285
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 3322
; LENGTH: 275
; TYPE: DNA
; ORGANISM: Cricetulus longicaudatus
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (172)..(192)
; OTHER INFORMATION: n is a, c, g, or t
US-11-128-049-3322

Query Match          64.8%; Score 16.2; DB 8; Length 275;
Best Local Similarity 85.7%; Pred. No. 1.9e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 5 AAGAGCGCAGACTGGTCTTCC 25
   ||||| ||||| ||||| |||||
Db 198 AAGAGCACTGACTGCTCTTCC 218

RESULT 45
US-11-128-049-6964
; Sequence 6964, Application US/11128049
; Publication No. US20060010513A1
; GENERAL INFORMATION:
; APPLICANT: Melville, Mark W.
; APPLICANT: Charlebois, Timothy S.
; APPLICANT: Mounts, William M.
; APPLICANT: Hann, Louane E.
```


APPLICANT: Sinacore, Martin S.
APPLICANT: Leonard, Mark W.
APPLICANT: Brown, Eugene L.
APPLICANT: Miller, Christopher P.
TITLE OF INVENTION: OLIGONUCLEOTIDE ARRAYS TO MONITOR GENE EXPRESSION AND METHODS FOR
FILE REFERENCE: 01997.027700
CURRENT APPLICATION NUMBER: US/11/128,049
CURRENT FILING DATE: 2005-05-11
PRIOR APPLICATION NUMBER: US 60/570,425
PRIOR FILING DATE: 2004-05-11
NUMBER OF SEQ ID NOS: 7285
SOFTWARE: PatentIn version 3.3
SEQ ID NO 6964
LENGTH: 275
TYPE: DNA
ORGANISM: Cricetulus longicaudatus
FEATURE:
NAME/KEY: misc feature
LOCATION: (172)..(192)
OTHER INFORMATION: n is a, c, g, or t
US-11-128-049-6964

Query Match 64.8%; Score 16.2; DB 8; Length 275;
Best Local Similarity 85.7%; Pred. No. 1.9e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 5 AAGAGCGCAGACTGCTCTTCC 25
||||| ||||| ||||| |||||
Db 198 AAGAGCACTGACTGCTCTTCC 218

RESULT 46
US-11-128-061-4682
Sequence 4682, Application US/11128061
Publication No. US20060003958A1
GENERAL INFORMATION:
APPLICANT: Melville, Mark W.
APPLICANT: Charlebois, Timothy S.
APPLICANT: Mounts, William M.
APPLICANT: Hann, Louane E.
APPLICANT: Sinacore, Martin S.
APPLICANT: Leonard, Mark W.
APPLICANT: Brown, Eugene L.
APPLICANT: Miller, Christopher P.
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES RELATED TO OLIGONUCLEOTIDE ARRAYS
FILE REFERENCE: 01997.027701
CURRENT APPLICATION NUMBER: US/11/128,061
CURRENT FILING DATE: 2005-05-11
PRIOR APPLICATION NUMBER: US 60/570,425
PRIOR FILING DATE: 2004-05-11
NUMBER OF SEQ ID NOS: 7285
SOFTWARE: PatentIn version 3.3
SEQ ID NO 4682
LENGTH: 425
TYPE: DNA
ORGANISM: Cricetulus griseus
FEATURE:
NAME/KEY: misc feature
LOCATION: (370)..(390)
OTHER INFORMATION: n is a, c, g, or t
US-11-128-061-4682

Query Match 64.8%; Score 16.2; DB 8; Length 425;
Best Local Similarity 85.7%; Pred. No. 2.1e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 5 AAGAGCGCAGACTGCTCTTCC 25
||||| ||||| ||||| |||||
Db 396 AAGAGCACTGACTGCTCTTCC 416

RESULT 47
US-11-128-049-4682
Sequence 4682, Application US/11128049
Publication No. US20060010513A1
GENERAL INFORMATION:
APPLICANT: Melville, Mark W.
APPLICANT: Charlebois, Timothy S.
APPLICANT: Mounts, William M.
APPLICANT: Hann, Louane E.
APPLICANT: Sinacore, Martin S.
APPLICANT: Leonard, Mark W.
APPLICANT: Brown, Eugene L.
APPLICANT: Miller, Christopher P.
TITLE OF INVENTION: OLIGONUCLEOTIDE ARRAYS TO MONITOR GENE EXPRESSION AND METHODS FO
FILE REFERENCE: 01997.027700
CURRENT APPLICATION NUMBER: US/11/128,049
CURRENT FILING DATE: 2005-05-11
PRIOR APPLICATION NUMBER: US 60/570,425
PRIOR FILING DATE: 2004-05-11
NUMBER OF SEQ ID NOS: 7285
SOFTWARE: PatentIn version 3.3
SEQ ID NO 4682
LENGTH: 425
TYPE: DNA
ORGANISM: Cricetulus griseus
FEATURE:
NAME/KEY: misc feature
LOCATION: (370)..(390)
OTHER INFORMATION: n is a, c, g, or t
US-11-128-049-4682

Query Match 64.8%; Score 16.2; DB 8; Length 425;
Best Local Similarity 85.7%; Pred. No. 2.1e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 5 AAGAGCGCAGACTGCTCTTCC 25
||||| ||||| ||||| |||||
Db 396 AAGAGCACTGACTGCTCTTCC 416

RESULT 48
US-11-108-172-584/c
Sequence 584, Application US/11108172
Publication No. US20050260177A1
GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun
APPLICANT: Lodes, Michael J.
APPLICANT: Secrist, Heather
APPLICANT: Benson, Darin R.
APPLICANT: Meagher, Madeleine Joy
APPLICANT: Stolk, John A.
APPLICANT: Wang, Tongtong
APPLICANT: Jiang, Yuqiu
APPLICANT: Smith, Carole L.
APPLICANT: King, Gordon E.
APPLICANT: Wang, Aijun
APPLICANT: Clapper, Jonathan D.
APPLICANT: Skeiky, Yasir A. W.
APPLICANT: Fanger, Gary R.
APPLICANT: Vedvick Thomas S.
APPLICANT: Carter, Darrick
TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
TITLE OF INVENTION: OF COLON CANCER AND METHODS FOR THEIR USE
FILE REFERENCE: 210121.471C15
CURRENT APPLICATION NUMBER: US/11/108,172
CURRENT FILING DATE: 2005-04-15
PRIOR APPLICATION NUMBER: US 10/025,380
PRIOR FILING DATE: 2001-12-19
PRIOR APPLICATION NUMBER: US 09/922,217
PRIOR FILING DATE: 2001-08-03
PRIOR APPLICATION NUMBER: US 09/833,263
PRIOR FILING DATE: 2001-04-10

```
; PRIOR APPLICATION NUMBER: US 09/649,811
; PRIOR FILING DATE: 2000-08-28
; PRIOR APPLICATION NUMBER: US 09/609,448
; PRIOR FILING DATE: 2000-06-29
; PRIOR APPLICATION NUMBER: US 09/575,251
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: US 09/519,444
; PRIOR FILING DATE: 2000-03-06
; PRIOR APPLICATION NUMBER: US 09/504,629
; PRIOR FILING DATE: 2000-02-15
; PRIOR APPLICATION NUMBER: US 09/480,321
; PRIOR FILING DATE: 2000-01-10
; PRIOR APPLICATION NUMBER: US 09/476,296
; PRIOR FILING DATE: 1999-12-30
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1130
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 584
; LENGTH: 431
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-11-108-172-584

Query Match      64.8%; Score 16.2; DB 8; Length 431;
Best Local Similarity 85.7%; Pred. No. 2.1e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 5 AAGAGCGCAGACTGCTCTTCC 25
   ||||| ||||| |||||
Db 122 AAGAGCACTGACTGCTCTTCC 102

RESULT 49
US-11-128-061-1040
; Sequence 1040, Application US/11128061
; Publication No. US20060003958A1
; GENERAL INFORMATION:
; APPLICANT: Melville, Mark W.
; APPLICANT: Charlebois, Timothy S.
; APPLICANT: Mounts, William M.
; APPLICANT: Hann, Louane E.
; APPLICANT: Sinacore, Martin S.
; APPLICANT: Leonard, Mark W.
; APPLICANT: Brown, Eugene L.
; APPLICANT: Miller, Christopher P.
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES RELATED TO OLIGONUCLEOTIDE ARRAYS
; FILE REFERENCE: 01997.027701
; CURRENT APPLICATION NUMBER: US 60/570,425
; PRIOR FILING DATE: 2004-05-11
; NUMBER OF SEQ ID NOS: 7285
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1040
; LENGTH: 469
; TYPE: DNA
; ORGANISM: Cricetulus griseus
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (370)..(390)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (426)..(450)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (453)..(465)
; OTHER INFORMATION: n is a, c, g, or t
; US-11-128-061-1040

Query Match      64.8%; Score 16.2; DB 8; Length 469;
Best Local Similarity 85.7%; Pred. No. 2.1e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 5 AAGAGCGCAGACTGCTCTTCC 25
   ||||| ||||| |||||
Db 396 AAGAGCACTGACTGCTCTTCC 416

Search completed: February 3, 2006, 16:20:24
Job time : 346.111 secs
```